UNIVERSITAT DE BARCELONA

Essays in Education, Fertility, and the Welfare State

Niclas Frederic Poitiers

PhD in Economics

Essays on Education, Fertility, and the Welfare State

Niclas Frederic Poitiers

2019

PhD in Economics | Niclas Frederic Poitiers

UNIVERSITAT DE BARCELONA

PhD in Economics

Thesis title:

Essays on Education, Fertility, and the Welfare State

PhD student:
Niclas Frederic Poitiers

Advisors:
Xavier Raurich
Concepció Patxot

Date:
May 2019

UNIVERSITAT DE BARCELONA

Für Regine, Jens & Philipp

Acknowledgements

Writing a thesis is a long and difficult journey, and one that I could not have hoped to conclude without the help of many friends and mentors along the way.

First, I want to express my deep gratitude to my advisers Xavier Raurich and Concepció Patxot. Xavier has helped me to learn how to develop models, write economic papers and develop the second chapter of this thesis. As much as I feared getting my manuscript back red with corrections, as much it helped me to become a better economist. Ció has supported me all the way along the third and forth chapter. I might not have continued the Ph. D. if it were not for her accepting me as a student and supporting me financially and she has helped me enormously in developing the third and forth chapter of this thesis. I am deeply in debt to her for that.

My coauthor and fellow student Gianko Michailidis has helped me not just in developing our joint paper, but also in managing the stress and anxieties of the Job Market. I really enjoyed working with him and I am very happy to have shared this part of the journey with him.

I want to thank the Universtiat de Barcelona School of Economics for giving me the opportunity of studying for a Ph. D. and supporting me in this studies. In particular, I want to thank Jordi Roca Solanelles for his support way beyond what one could expect of him. Especially in the beginning I would have been lost in the forest of Spanish and Catalan bureaucracy if it were not for him. I want to thank as well Elisabet Viladecans-Marsal for her support and her dedication to the program. I am very proud to have studied at this institution and to have seen it develop throughout the four years I have been there.

I am grateful to the Universitat de Barcelona for granting me a scholarship that allowed me to dedicate myself to the research that cumulated in this thesis. Furthermore, I want to thank the Societat Econòmica Barcelonesa d'Amics del País for granting me the scholarship that enabled my research stay at Northwestern University.

Throughout the development of the thesis I got help and advice from experienced researchers, to whom I am very thankful. I want to thank Aydan Dogan, Ester Manna, Gemma Abío, Marc Teignier, Montserrat Vilalta-Bufí, Vahagn Jerbashian, Joseph Zeira, and many others for their feedback and advice. I am very grateful to Matthias Doepke, who has welcomed me as a visiting researcher at

Northwestern University and supported me in the Job Market.

Many of my fellow students have helped me throughout the journey. The ideas that I develop in this thesis benefited enormously from discussing it with the people with whom I shared offices. Athina Raftopoulou, Bernard Moscoso, Cesar Blanco, Cynthia Armas, Diego Ocampo, Ferran Armada, Giorgos Papadomichelakis, Manish Singh, Natassa Papadopoulou, and Till Hollstein made the everyday live as a pre-doctoral researcher a very nice experience. I want to thank Nicola Rubino for helping me out when I was lost in econometrics, and Tania Fernández-Navia who helped me very much in sharpening my ideas through discussions as well as managing the stress of life as a Ph. D. student. Kinga Tchorzewska supported me substantially through the stressful journey in the Job Market, and I am very grateful to her for that. There are many more that I would like to thank than I can list here. Working on a Ph. D. can sometimes feel hard and isolating, but the community at Universitat de Barcelona School of Economics helped me very much to deal with that.

I also want to thank Danisz Okulicz and Christoph Gschnaidtnerr for their companionship in Chicago, which made the research stay a very nice experience. A particular thank goes to my flatmates Raffaella Freyre and Alba Asensio, who have been like a family to me here in Barcelona. I am also deeply grateful to Nicola Heins and Heike Struß for their support throughout my life and in pursuing my studies.

Finally, I want to thank my family. Regine, Jens, and Philipp have always supported me and lifted me up in difficult times. They believed in me when I did not myself. This thesis is dedicated to them.

Contents

List of Tables

List of Figures

1 Introduction

In countries in the developed world, income inequality is increasing, while technological and societal changes open labour market opportunities for women. At the same time they are undergoing an important demographical transition with decreasing fertility and increasing population ageing. All these trends affect the decisions that different generations make over the life-cycle. In this thesis, I investigate the role that these trends play for education, fertility, and pensions using overlapping generations models. This type of economic models allow us to investigate the decisions of economic agents along the life cycle taking into account the interaction with other generations.

The structure of this thesis will follow the life-cycle of an economic agent: In the second chapter I investigate how income inequality is affecting the educational decisions of young agents and their social mobility. In the third chapter, I investigate the decision of couples to have children and educate them and how this decision is affected by changes in the gender wage gap. In the forth chapter, I investigate how in an economy with a pay-as-you-go pensions system the intragenerational conflict between rich and poor as well as the intergenerational conflict between old and young about the allocation of public resources into education and pensions are affected by income inequality and population ageing. The research that I conduct in this thesis aims at contributing to our understanding of the mechanisms at play.

In the second chapter of this thesis, I investigate how income inequality is affecting the education attainment of the young. The education that we receive when young strongly influences our earning ability and position in the income distribution over our life time. Thus it directly affects the extend to which our income is related to that of our parents, our intergenerational social mobility. The empirical literature on intergenerational social mobility shows that countries with higher income inequality have lower levels of intergenerational social mobility. In countries with high income inequality, as the U. S., the income of children depend stronger on the one of their parents than in countries with low income inequality like Denmark. This negative association between income inequality and intergenerational social mobility is called the "Great Gatsby Curve". Corak (2013a) shows such a negative relationship across countries, and Chetty et al. (2014b) also finds it within the U. S.

Understanding the mechanism by which income inequality decreases intergenerational social mobility is of increasing importance, as income inequality is increasing (Piketty, 2013) and the size of the middle class is decreasing (Acemoglu and Autor, 2011). Social mobility influences economic growth, as low levels of intergenerational social mobility imply that we are not allowing all individuals to achieve their full potential. Galor and Zeira (1993) argue that it is in fact the opportunity to attain higher education for the whole population that is driving economic growth of developed countries.

Most of the literature on social mobility focuses on college education and credit constraints on parents in acquiring higher education. In Galor and Zeira (1993), Galor and Zang (1997), and Alonso-Carrera et al. (2012), income inequality is affecting social mobility through the share of parents that are rich enough to provide higher education for their children. In Hassler and Rodríguez Mora (2000) and Hassler et al. (2007) the two are related through the speed of technological progress and the returns to education and effort. Higher rates of technological growth lead to higher income inequality but also higher levels of social mobility. In Piketty (1995) and Checchi et al. (1999) persistence of income across generations is driven by beliefs about ability and effort formed through the histrocial experices of families.

Jäntti et al. (2006) show that differences between countries in the income mobility over generations are driven by the downward mobility of children from high income families, and the upward mobility of children from poor backgrounds. In this chapter I will focus on the latter and investigate how education and hence the upward mobility of poor children is affected by income inequality.

In particular, I investigate the effect that wage premiums has on high school attainment of poor children. High school education plays a very important role in explaining the mobility patterns of children from low income households. Restuccia and Urrutia (2004) show that even though college education accounts for most differences in incomes, it is the investment into early education that explain the persistence of income over generations.

I develop an overlapping generations model based on Galor and Zeira (1993), where children decided which level of education to attain based on a transfer that they receive from their parents. This model contributes to the literature by directly modelling high school dropout and by allowing young students to choose higher education levels even when facing a binding credit constraint. In the model, poor students face a trade-off between studying or working, while this opportunity cost is not relevant for students from high income families.

I show that in the model an increase in the opportunity cost of high school education in the sense of an increase in the wage of young high school dropouts relative to the wage of young high school graduates would decrease the high school graduation rate. An increase in the returns on high school education would increase the graduation rate. I find some empirical evidence that this kind of opportunity cost of education is relatively more important than the returns to education. Motivated by this findings, I show in the model that a policy decreasing the opportunity cost of high school education by decreasing taxes on young high school graduates financed through a higher tax on older high school graduates (which decreases the returns on education) would increase overall graduation rates. In the model such a policy applied to college education could increase college graduation rates and decrease the share of college students facing a binding credit constraint when going to college. This could potentially decrease the levels of student debt.

In the third chapter, I focus on the decision of adults to have children. There was a strong decline in fertility during the 20th century in Germany. Households have less and less children, and the completed fertility rate (the number of children a woman has over her lifetime) has dropped below the rate that would be necessary to sustain the population without migration. This has important implications for the labour market and puts strains on the welfare state (Abío et al., 2004).

Generally, the decline in fertility is understood as the result of a substitution between the number of children and the investment into their education, the so called "quantity/quality" trade-off (Becker, 1960). As families become richer and education more important, they choose to have fewer children in order to provide higher education to them. In Galor and Weil (1996) this is driven by a change in the returns on mental labour relative to physical labour caused by capital accumulation. This leads to a decrease in the gender wage gap and makes children more costly. In Galor and Moav (2000), de la Croix and Doepke (2003), Galor and Tsiddon (1997), the quantity/quality trade-off has implications on growth as it interacts with human capital accumulation. In richer countries, families have less children and invest more into their human capital. Moreover, income inequality has effects on economic growth as it leads to differences in fertility and education between rich and poor.

Recently, there is an increasing interest in the role that childlessness plays for the demographic transition. Gobbi (2013) shows that childlessness is negatively associated with the fertility rate, and Baudin et al. (2015) and Baudin et al. (2018)

shows its importance for understanding fertility patterns within the U. S. and for developing countries, respectively. The quantity/quality trade off cannot account for childlessness, as in order to substitute the number of children for investment into their education, there has to be at least one child into who's education a family can invest (Aaronson and Mazumder, 2008). In the literature on childlessness, the decisions not to have children is the outcome of preference transmission over generations (Gobbi, 2013), or the interaction between social fertility of the poor and "opportunity" driven childlessness of the rich (Baudin et al., 2015, 2018). However, this literature does not take investment into education into account, the main driver of the quantity/quality trade-off.

I find in an empirical decomposition of the decline in the fertility rate in Germany, that although until the generation born in the 1950s, the entire decline in fertility was due to a decrease in the number of children per women that has at least one child (with low and stable rates of childless women), for the generations of women born after 1950 the decline in fertility was entirely due to an increase in the share of childless women.

The third chapter of my thesis aims at reconciling the literature on childlessness with the one on the quantity/quality trade-off. I am investigating the the opportunity cost of childbirth for women in terms of labour market outcomes to understand the interaction between fertility and childlessness. Kleven et al. (2018) show that there is a "child penalty" for women in terms of a decline in the income after childbirth. Men do not show any change in their income after the birth of their first child.

I conduct an event study analysis based on Kleven et al. (2018) in order to asses the child penalty for Germany. I find in this analysis that the decline in labour income of women after childbirth is driven entirely by a decline in labour force participation and hours worked. The labour force participation as well as hours worked (conditional on participating on the labour market) decline strongly, whereas there is no significant change in the hourly wage. Thus women face a large opportunity cost of having children in terms of work time. Moreover, this child penalty is independent of education.

Based on these empirical findings, I develop an overlapping generations model with quantity/quality trade off and a child penalty in terms of women's time. I show in a calibration exercise, that this model can generate both the decline in number of children for women with at least one child and the increase in childlessness through a decrease in the gender wage gap that is matched to data on the female labour force participation rate. This analysis shows that both childlessness as well as the opportunity cost of childbirth for women in terms of work time are very important in order to understand the demographical change in developed countries.

The forth chapter of my thesis is a joint work with Gianko Michailidis on the effect of population ageing and income inequality on public education and public pensions. Parallel to the strong increase in income inequality in developed countries, caused by the decreases in the fertility rate and increases in life expectancy, the population of developed countries are ageing at an increasing pace.

This has an effect on public welfare spending, as it increases the number of people receiving public pensions, as well as it changes the demographics of the electorate. Old people have less incentive to spend on education as they will not benefit from the future reward of investments in human capital (Poterba, 1997). Therefore the increase in population ageing leads to an intensification of the intergenerational conflict over public spending on education and pensions. At the same time, the increase in income inequality has implications for the welfare state as well. It increases the intragenerational conflict between rich and poor over redistribution in the form of public education spending.

The literature on the political economy of education and pensions has treated these two conflicts in separation. Browning (1975) and Ono and Uchida (2016) investigate the intergenerational conflict looking at the effect of population ageing on public expenditures on education and pensions. Browning (1975) shows that the increasing political power of the elderly shifts resources towards pensions. In Ono and Uchida (2016), an increase in longevity increases total public pensions spending, but has a non-linear effect on education. In the literature on the intragenerational conflict about public education, Stiglitz (1974) shows that in political economy models the educational outcomes depend on whether education is perceived as public or private good. In de la Croix and Doepke (2009) an increase in income inequality decreases participation in public education, which leads to an increase in education spending per student. Levy (2005) develops a political economy model of education with endogenous party formation, where the effect of income inequality depends on whether young or old are in the majority.

We contribute with this study to the literature by considering the intergenerational conflict between young and old and the intragenerational conflict between rich an poor jointly. This allows us to investigate the interaction between them. We develop an overlapping generations model based on de la Croix and Doepke

(2009) with a public pay-as-you-go pensions system and public and private education.

In this model, an increase in income inequality decreases participation in public education. Public education has to be provided for less students, and thus there are resources to increase per student education spending, decrease taxes, and increase pensions. An increase in population ageing has the opposite effect. More public funds are needed to finance pensions, leading to a decrease in per student public education spending and pensions, and an increase in taxes.

Using panel data on OECD countries, we find evidence for a negative effect of population ageing on education spending per student, but only mixed results regarding the effect of income inequality on education spending per student.

2 The Impact of Wage Differentials on Intergenerational Social Mobility

2.1. Introduction

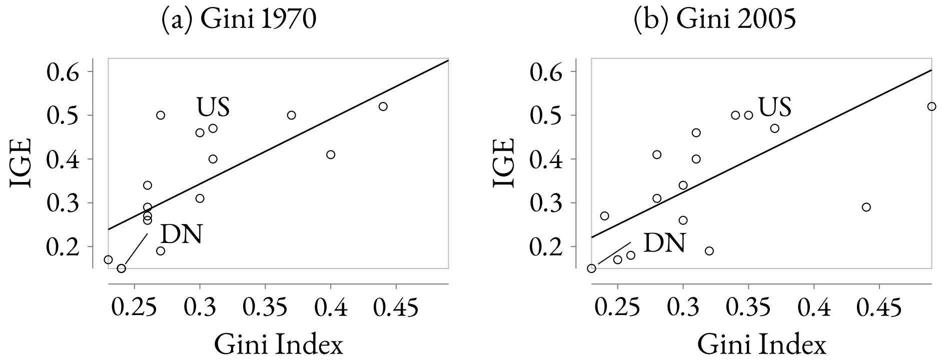
There is a strong correlation in cross country data between income inequality and intergenerational social mobility. Empirical studies suggest that this is driven by the extent in which children with parents in the highest and lowest income quintiles are downward and upward mobile over generations, respectively. One of the most important drivers of social mobility is education. I investigate in this paper to which extent income inequality, measured in wage differences, is driving differences in intergenerational social mobility. This is of particular interest as income inequality and employment polarisation are increasing in developed countries. In the literature on intergenerational mobility, inequality hinders intergenerational mobility mainly by imposing binding credit constraints on the ability of parents to acquire education for their children. The parents decide how much education to buy form their children. The ability to buy education is constrained by the budget of the parents. In this paper I explore a different approach: Not the parents but the children themselves decide on how much education to acquire They decide whether to educate or to work and earn an unskilled wage. Children from poor backgrounds can still choose to educate, but since they receive less financial support from their parents, the value of the outside option to work is much more important for them. This approach allows me to identify the effects of education finance policies on the educational outcomes of children from poorer and richer family backgrounds. The model as well as the empirical exercise suggest that opportunity cost to education play a very important role in explaining education outcomes. They are relatively more important than the returns to education. This has important implications for the effect of wage polarisation on education decisions. The second contribution of this paper to the existing literature is that it directly models high school drop out. High school drop out is one of the main factors explaining differences in social mobility. Most models so far focus on the role of college education. Evidence in the empirical literature suggests though that high school dropout plays a more important role in explaining persistence of education levels.

The degree of income inequality in a country is negatively associated with intergenerational mobility. The more unequal a country, the more persistent is income over generations. This is called the “Great Gatsby Curve”. One measure of intergenerational social mobility is the intergenerational earnings elasticity (IGE). It is the elasticity between a child’s and their parents’ income. A higher IGE means that children’s incomes depend stronger on their parents’ income, i. e. lower intergenerational income mobility. Corak (2013a,b) estimates the IGE for a variety of countries. In Table 2.1, I display his IGE estimations and a variety of indices of income inequality (including some based on transition matrices of Jäntti et al., 2006). The correlation of the IGE with inequality indices are shown in Table 2.2.

The IGE features a high positive correlation with all of them. As a higher IGE implies less intergenerational income mobility, this implies a negative relationship between income inequality and intergenerational income mobility. The 1970 Gini-index and 2005 Gini-index have correlations of 0.69 and 0.71 with the IGE, respectively. The IGE and Gini-index relation is depicted in Figure 2.1. One can see the strong negative association of inequality and social mobility, the “Great Gatsby Curve”. Chetty et al. (2014a) find such a relation also within the U. S. Areas with a lower share of middle-class residents have lower levels of social mobility. Thus regions with higher levels of income inequality tend to have lower levels of intergenerational income mobility.

In order to better understand this relationship, it is helpful to look at transition matrices, which provide a more detailed picture than IGE estimates. For each income quantile, they give the probabilities for a child born into a family in this quantile to end up in each income quantile. Jäntti et al. (2006) compares the transition matrices of Scandinavian countries with the ones of the U. K. and the U. S. Jäntti et al. (2006) show that the most important dimension in which countries with high intergenerational mobility (Scandinavian countries) differ from ones with low intergenerational mobility (U. S. and U. K.) is the persistence of high and low income families (top and bottom 20%) in their quintiles. Figure 2.2 shows a representation of the transition matrices of Denmark and the U. S. estimated by Jäntti et al. (2006). Each line represents the probabilities for a child from one in-

Figure 2.1.: *The Great Gatsby Curve*



NOTE: IGE values from Corak (2013b) and mean Gini index estimates from Standardized World Income Inequality Database (SWIID), see Table 2.1. The lines represent fitted values. The values for the U. S. and Denmark are highlighted, of which the transition matrices are compared in Table 2.2.

come quintile of ending up in each income quintile. One can see that in the U. S., high income families have a lower downward mobility and low income families have a lower upward mobility than in Denmark (see also Table 2.3).

The view that the bottom and the top of the income distribution are responsible for the relationship between income inequality and social mobility is supported by Corak et al. (2014). They observe that the main differences in absolute earnings-mobility (a son's income relative to his father's) between the U. S., Canada, and Sweden are in the extent of downward mobility (sons earning less than their fathers) from the top of the income distribution. Furthermore, Couch and Lillard (2004) observe non-linear patterns in income persistence for the U. S. and Germany.[1] They find evidence that earnings are more persistent over generations for high income families than for those with lower income.

Educational attainment determines to a large extent one's lifetime income level.[2] It is also strongly dependent on parental background. In the U. S., children from affluent families have a much lower probability of dropping out of high school. In Table 2.4, I show the educational attainment by parental income in the U. S. using the National Longitudinal Survey of Youth 1979 (NLSY-79). One can see that the probability of dropping out of high school is very low for children of affluent families (0.08), while it is very high for children from the lowest income quintile (0.36). This resembles the pattern observed in the transition matrices. The probabilities to end up in the lowest income quintile is very similar: 0.07 for children of high in-

[1]Earnings of sons are stronger related to those of their fathers when the father is richer.
[2]For a discussion of U. S. wage premiums see Lemieux (2006)

Table 2.1.: *Measures of Inequality and Mobility*

	Gini 1970	Gini 2005	90/10	90/50	50/10	IGE	I_λ	I_{Trace}	I_{Cross}
Denmark	0.24	0.23	2.8	1.5	1.8	0.15	0.81	0.93	0.35
Finland	0.25	0.26	3.2	1.7	1.9	0.18	0.80	0.93	0.35
Norway	0.23	0.25	3.1	1.6	1.9	0.17	0.78	0.92	0.34
Sweden	0.26	0.24	3.5	1.7	2.1	0.27	0.78	0.92	0.34
United Kingdom	0.27	0.35	4.4	2.0	2.2	0.50	0.79	0.93	0.34
United States	0.31	0.37	5.8	2.2	2.7	0.47	0.66	0.87	0.30
Australia	0.26	0.30	4.4	1.9	2.3	0.26			
Canada	0.27	0.32	4.5	1.9	2.3	0.19			
Chile	0.44[1]	0.49	9.0	3.3	2.7	0.52			
France	0.40	0.28	3.5	1.9	1.9	0.41			
Germany	0.30	0.28	3.7	1.8	2.0	0.31			
Italy	0.37	0.34	4.2	1.9	2.2	0.50			
Japan	0.26	0.30	5.3	2.0	2.7	0.34			
New Zealand	0.26	0.33	4.3	1.9	2.2	0.29			
Spain	0.31	0.31	4.8	2.0	2.5	0.40			
Switzerland	0.30[2]	0.31	3.4	1.8	1.9	0.46			

NOTE: Gini index estimates are the mean estimates from the SWIID. Other measures of inequality are form the OECD, based on the year 2009, except for Australia with values for the year 2008. IGE estimates are from Corak (2006). 90/10, 90/50 and 50/10 are the ratios between the corresponding income percentiles. Measures of mobility are based on transition matrices based on the age-corrected transition matrices of father and sons from Jäntti et al. (2006). The indices based on the transition matrices are the following. The first index is based on the second largest eigenvalue λ_2 of the mobility matrix: $I_\lambda = 1 - |\lambda_2|$. In the following m is the number of rows of the matrix and p_{ij} is the transition probability from quantile i to quantile j. π_i is the long run probability of being in quantile i (i. e. $\frac{1}{m}$). The second index of mobility in transitions matrices is based on the trace of the matrix: $I_{\text{Trace}} = \frac{m-\sum_{i=1}^{m} p_{ii}}{m-1}$. The third index is based on the expected number of income brackets crossed: $I_{\text{Cross}} = \frac{1}{m-1}\sum_{i=1}^{m}\sum_{j=1}^{m}\pi_i p_{ij}|i-j|$. For all three transition matrix indices, higher values indicate lower levels of intergenerational mobility.

[1] 1968

[2] 1971

come families and 0.4 for children of low income families. As in the transmission of income, the difference between the top and the bottom of the income distribution is especially strong. That raises the question of how educational decisions at the top and the bottom of the income distribution are influenced by income inequality. I want to address this question by looking at wage premiums and the opportunity cost of education.

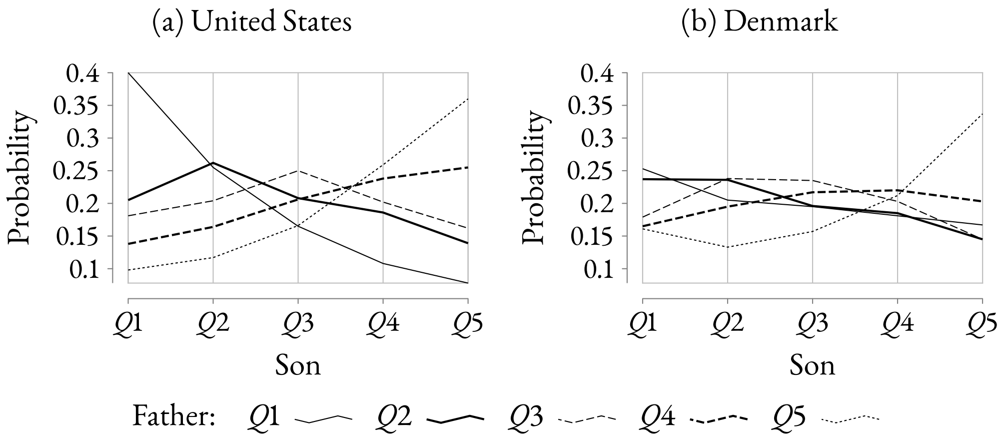
It seems to be of particular importance to understand the role that wage premiums play for intergenerational mobility, since income inequality has increased over the last decades (see Piketty, 2013) and this increase has taken the form of a polarisation of labour markets. In the last two decades there was a polarisation of the labour market, with less middle wage jobs and more low and high wage jobs. Between 1993 and 2010 the share of hours worked in middle wage occupations declined by 8 percentage points in Europe and by 6 percentage points in the U. S. (Goos et al., 2009; Acemoglu and Autor, 2011). Acemoglu and Autor (2011)

Table 2.2.: *Correlations of IGE and Inequality Measures*

	Gini 1970	Gini 2005	90/10	90/50	50/10	IGE
Gini 1970	1.00	0.67	0.61	0.72	0.31	0.69
Gini 2005	0.67	1.00	0.92	0.94	0.71	0.71
90/10	0.61	0.92	1.00	0.97	0.84	0.56
90/50	0.72	0.94	0.97	1.00	0.71	0.62
50/10	0.31	0.71	0.84	0.71	1.00	0.47
IGE	0.69	0.71	0.56	0.62	0.47	1.00

NOTE: Gini index estimates are the mean estimates from the SWIID (see Table 2.1). Other measures of inequality are from the OECD, based on the year 2009, except for Australia with values for the year 2008. IGE estimates are taken from Corak (2006).

Figure 2.2.: *Transition Probabilities*



NOTE: Probability of a son with a father in income quintile Q_i to be in income quintile Q_j. The probabilities are from Jäntti et al. (2006), see Table 2.3.

found for the U. S. that during the same period, the wage growth in middle wage occupations has lagged considerably behind the wage growth in low and high wage occupations. This implies not only a polarisation of the labour market, but also of wages. Deschênes (2001), Lemieux (2006) and Acemoglu and Autor (2011) observe a convexificiation of the premiums on education since 1980, i. e. a strong increase in the return on higher education. There is no evidence of a polarisation of wages in Europe yet. There is an increase of upper tail inequality, but no decrease of lower tail inequality in the U. K. and Germany (Manning et al., 2007; Antonczyk et al., 2010).[3] Yet, analogous to the convexification of educational returns in the U. S., Pereira and Budría (2005) and Lindley and Machin (2011) showed that

[3]Here, upper tail inequality relates to the ration of the 90th to the 50th income percentile and lower tail inequality to the ration of the 50th to the 10th income percentile.

Table 2.3.: *Transition Matrices*

		Denmark					U. K.				
		Son					Son				
		Q_1	Q_2	Q_3	Q_4	Q_5	Q_1	Q_2	Q_3	Q_4	Q_5
Father	Q_1	0.253	0.205	0.195	0.181	0.167	0.303	0.235	0.165	0.174	0.122
	Q_2	0.237	0.236	0.196	0.185	0.145	0.241	0.227	0.182	0.193	0.157
	Q_3	0.179	0.238	0.235	0.203	0.145	0.188	0.195	0.227	0.206	0.184
	Q_4	0.165	0.195	0.217	0.220	0.203	0.161	0.175	0.229	0.195	0.240
	Q_5	0.161	0.133	0.157	0.212	0.337	0.107	0.168	0.197	0.231	0.297

		U. S. NLSY-79							U. S. Chetty				
		Son							Child				
		Q_1	Q_2	Q_3	Q_4	Q_5			Q_1	Q_2	Q_3	Q_4	Q_5
Father	Q_1	0.400	0.254	0.165	0.108	0.074	Parents	Q_1	0.337	0.242	0.178	0.134	0.109
	Q_2	0.205	0.262	0.208	0.186	0.139		Q_2	0.280	0.242	0.198	0.160	0.119
	Q_3	0.181	0.204	0.250	0.202	0.162		Q_3	0.184	0.217	0.221	0.209	0.170
	Q_4	0.138	0.164	0.206	0.238	0.255		Q_4	0.123	0.176	0.220	0.244	0.236
	Q_5	0.098	0.117	0.166	0.259	0.360		Q_5	0.075	0.123	0.183	0.254	0.365

NOTE: Transition matrices from Jäntti et al. (2006), income quintile transition matrices for sons and fathers, corrected for age. U. S. data from Chetty et al. (2014a) generally linking parents to children (both sons and daughters).

the inequality in post-graduate wages in the EU has increased.

There is not only evidence of a polarisation of employment and education premiums, but also of increased polarisation of educational efforts. Putnam et al. (2012) find a growing gap between high school students from upper and middle class backgrounds with respect to participation in soft-skill building activities.[4] Ramey and Ramey (2010) observe a considerable increase in the time spent with children for middle and upper class parents since the mid-1990s, and Kornrich and Furstenberg (2013) show that the investment into children's education is increasingly unequal. Bailey and Dynarski (2011) find increasing dependence of college attendance on income for the period of 1961 to 1982, driven by an increase in the college attendance of daughters of high income families, but this has stabilised after 1982 (Chetty et al., 2014b). Lindley and Machin (2012) find for the U. K. that with increasing length of education, the importance of the family background is increasing.

Although the pattern of increased income inequality and convexification of the return on education point in the direction of less intergenerational social mobility, there is no agreement in the literature yet on whether intergenerational social mobility has actually decreased as a consequence. To the best of my knowledge, there are only attempts to estimate trends in social mobility for the U. S. Point in

[4]The gap in participation in extracurricular activities, i. e. sports and academic clubs, is increasing. These participations are a strong predictors of future success (Putnam et al., 2012).

time measurements indicate a strong increase of the IGE. Estimates of the IGE for the 1960s to 80s are around 0.2 (Becker and Tomes, 1986), Solon (1999) estimates it for the 1990s at around 0.4 and current estimates are around 0.6 as in Mazumder (2005).[5] Aaronson and Mazumder (2008) estimate the U. S. trend of the IGE based on Integrated Public Use Microdata Series (IPUMS). The child's year and state of birth are used to construct predicted parents' incomes. Their IGE estimations track the upward trend in income inequality between 1970 to 2000 very closely. Hertz (2007) and Lee and Solon (2008) estimate IGE trends based on the Panel Study of Income Dynamics (PSID) over the same time period. The estimations show no clear trend in the IGE.[6] Chetty et al. (2014b) estimate social mobility indices based on de-identified tax records and college attendance rates. They find no trend in social mobility measurements based on income rank, but as income inequality increases, the consequences of rank mobility has increased.[7]

In this paper, I develop an overlapping generations model (OLG model) based on Galor and Zeira (1993), with three levels of education where children choose their education level based on a transfer that they receive from their parents. The three levels of education reflect the different intergenerational mobility patterns of poor, middle class, and rich families observed in the empirical literature. Letting the children instead of the parents choose the education level allows to identify opportunity costs of education that are relevant for children form poor families but not for children from rich families. In particular, I consider the opportunity cost of education in terms of income that young adults can earn if they drop out of education and work instead. In Galor and Zeira (1993) and the related literature, parents choose the education level based on the costs of education for them. Parents facing a binding credit constraint cannot afford higher education for their children. In this paper children also face a credit constraint, but even if this credit constraint is binding they can choose a low level of consumption in the first part of their lives in order to acquire higher education and have a higher income in the future. Poor children cannot smooth consumption over both periods. The alternative of working instead of studying is therefore of higher importance for them

[5] A simple comparison of estimates with different samples and life-time income definitions suffers form comparability problems. For a discussion, see Hertz (2007).

[6] Lee and Solon (2008) observe a small increase for daughters. Hertz (2007) does estimations with four different specifications. In one he finds a positive trend in IGE, however, the other specifications show no trend at all.

[7] The probability of a child born into a family at the lowest quintile of the income distribution to reach the highest quintile is 8.4% for the 1971 cohort and 9% for the 1986 cohort.

than for rich children.

I find in the model that an increase in the opportunity cost decrease educational attainment by the poor, whereas an increase in the return on education increases educational attainment overall. Furthermore, when keeping the transfer from the parents unchanged, changes in the wages of graduates have an impact on the number of graduates not facing a binding credit constraint, whereas changes in the wages of individuals that drop out from education just affect the number of graduates facing a binding credit constraint. In an empirical assessment of the relative importance of the opportunity costs of education using U. S. data, I find that changes in the opportunity cost have a much stronger influence on the high school dropout rate of men than the return on education. Therefore, I propose a policy where the costs of education are paid by graduates in later stages of their life through taxes, which reduces the return on education but decreases the opportunity cost of education. This policy has the advantage that it does not imply transfers between educational groups, decreases income inequality due to age differences, and increases the number of graduates as well as the number of graduates not facing a binding credit constraint. Furthermore, this paper suggests that policies affecting the income distribution should not only be assessed in their effect on the overall level of income inequality, but also in how these policies affect the incentives for educational attainment, in particular the opportunity cost of education.

Current models of social mobility do mainly explain the impact of income inequality on social mobility through credit constraint agents (Galor and Zeira, 1993). Only rich agents with income or wealth above a certain threshold acquire education. Higher income inequality implies that more individuals are below the threshold for acquiring education or, as in the case of Moav and Galor (2004), can acquire an optimal level of education.[8] Alonso-Carrera et al. (2012) further develop the model of Galor and Zeira (1993) in order to allow for fiscal policies. In their model, using labour taxes instead of inheritance taxes increase human capital accumulation, while the impact of such a policy on income inequality would depend on the initial distribution of human capital. According to the analysis of Jäntti et al. (2006), it is not the middle class at the threshold of affording education that is responsible for lower mobility, but the upper and lower class, which in these mod-

[8] Galor and Moav (2006) argue that the increase in public education after the industrialisation was not the result of class struggles, but a result of an interest of the capitalist class in educated workers. In this model, the binding credit constraint is not overcome by a sufficient decrease of inequality, but by a political interest in taxation of the upper class.

els would always be above or below the income threshold.[9] Piketty (1995) and Checchi et al. (1999) model private and public investment into human capital as the results of beliefs about ability and effort. Differences in social mobility are the outcomes of differences in experienced mobility of dynasties. Both the theories of Piketty (1995) and Checchi et al. (1999) do imply that the differences in social mobility are a result of long term differences (over several generations) in countries' economic structures.

In Galor and Tsiddon (1997) and Hassler and Rodríguez Mora (2000), income inequality and social mobility depend on the rate of technological progress. In their models, technological progress increases inequality, as it increases the return on skill, but it also decreases the role of inherited human capital and thus increases social mobility. The evidence presented here goes the other way round: more income inequality is linked to less intergenerational social mobility, not more. Alonso-Carrera et al. (2016) use a framework similar to Galor and Tsiddon (1997) and Hassler and Rodríguez Mora (2000), but come to the opposite conclusion. They study the interaction between the education decision and the choice of occupations with different effort levels. An increase in the return on effort for high skilled decreases the frequency of high wages of low skilled, potentially increasing income inequality while reducing intergenerational social mobility. In Hassler et al. (2007) income inequality affects social mobility through two channels: through its effect on incentives for education and through its effect on parents ability to pay for their children's education. They emphasise the role of public education for mitigating the latter effect.

There are some studies on the reasons for high school dropout, but they are not considering social mobility. Eckstein and Wolpin (1999) study the causes of dropouts from high school using the NLSY-79. Students that drop out of high school have lower expectations about the reward of graduating.[10] McNeal (1997) studies the effects of employment during high school on the probability of dropping out of high school. The job type and the intensity have strong effects on the probability of dropping out of high school. The study most similar to my approach is Restuccia and Urrutia (2004). They developed an OLG model distinguishing between college and early (i. e. pre-college) education. They argue that

[9] And as Chetty et al. (2014a) point out, the middle class is actually the most mobile class, its lacking is impeding mobility strongly.

[10] Eckstein and Wolpin (1999) argue that a prohibition of working for high school students would have only limited impact, as the traits the children have when they come to school play an important role.

Table 2.4.: *High School Dropouts*

	Total	Q_1	Q_2	Q_3	Q_4	Q_5	Q_1	Q_2	Q_3	Q_4	Q_5
HS dropout[1]	1511	474	425	295	209	108	0.36	0.32	0.23	0.16	0.08
HS[2]	1906	408	397	412	392	297	0.31	0.30	0.31	0.29	0.22
HS+[3]	3200	436	501	603	731	929	0.33	0.38	0.46	0.55	0.70
Total	6617	1318	1323	1310	1332	1334	1.00	1.00	1.00	1.00	1.00

NOTE: Education level by parents' income quintiles. The data are from the NLSY-79 containing 14-22 year olds in 1979. Income quintiles correspond to the mean reported values before the age of 18.
[1] Highest attended grade <12.
[2] Highest grade attended 12.
[3] At least some college education.

about half of intergenerational income persistence is due to parental investments into early education, but college education accounts for most of the disparity.[11]

My approach contributes to this literature by studying the influence of wage differentials for educational choices. To this end, I introduce the decision of high school drop out into a Galor and Zeira (1993) framework. I developed an OLG model that captures the education decision of agents with respect to both high school and college. The agents make their choice regarding their individual opportunity cost of education and the return on education. In contrast to Restuccia and Urrutia (2004), I focus on the role of inequality by directly modelling the educational choice in pre-college education. As the evidence presented above indicates, this pre-college educational choice is crucial for understanding differences in social mobility. The second contribution of my model is that it distinguishes between the cost of education for parents and the incentive for children to graduate, i. e. the cost of education for the children in terms of utility. Children from low income families have a lower level of consumption as students, and dropping out of high school, i. e. giving up future income for current income, is a much more attractive option. On the other side of the income distribution, children from rich families already enjoy a high level of consumption during education and thus there is no incentive to drop out of high school.

The rest of the paper is structured in the following way: Section 2.2 introduces the model, Section 2.3 explains the educational choice of agents in the model, Section 2.4 explores the baseline comparative statics in the model, in Section 2.5 I make an empirical assessment of the relative importance of wage premiums, and in Section 2.6 I discusses the effectiveness of policies.

[11] Early education accounts for the largest part of the persistence, as younger parents are more strongly constrained in their budget for educational expenses.

2.2. The Model

I propose an OLG model based on Galor and Zeira (1993) of intergenerational social mobility in which agents decide which level of education to acquire. The model consists of three different levels of education in order to model the different mobility patterns observed for low, middle an high wage groups. This allows me to study the effect of wage polarisation on educational attainment. The education decision is defined by two forces: the opportunity cost of education and the return on education. The former are only relevant for children from poor backgrounds, whereas the latter also matter for rich children. Each individual lives for three periods. Each individual has one parent and one child, thus there is no population growth. The timing in the three periods of life is as follows: In the first period, the agent is born, he goes to school and can decide to drop out in order to earn the wage of unskilled workers for this period and the rest of his life. He can decide to finish school and work for the rest of his life for the high school graduate wage, or he can go to college, and does not earn anything in this period, but earns a college wage for the rest of his life. In the second period, the agent works and gets one child. He provides a transfer to his child. In the third period, the agent retires and consumes from his savings.

In the following, the index N denotes drop out of high school, H high school graduation and C college graduation. Depending on his education level, the agent has the following incomes in the first period of his life:

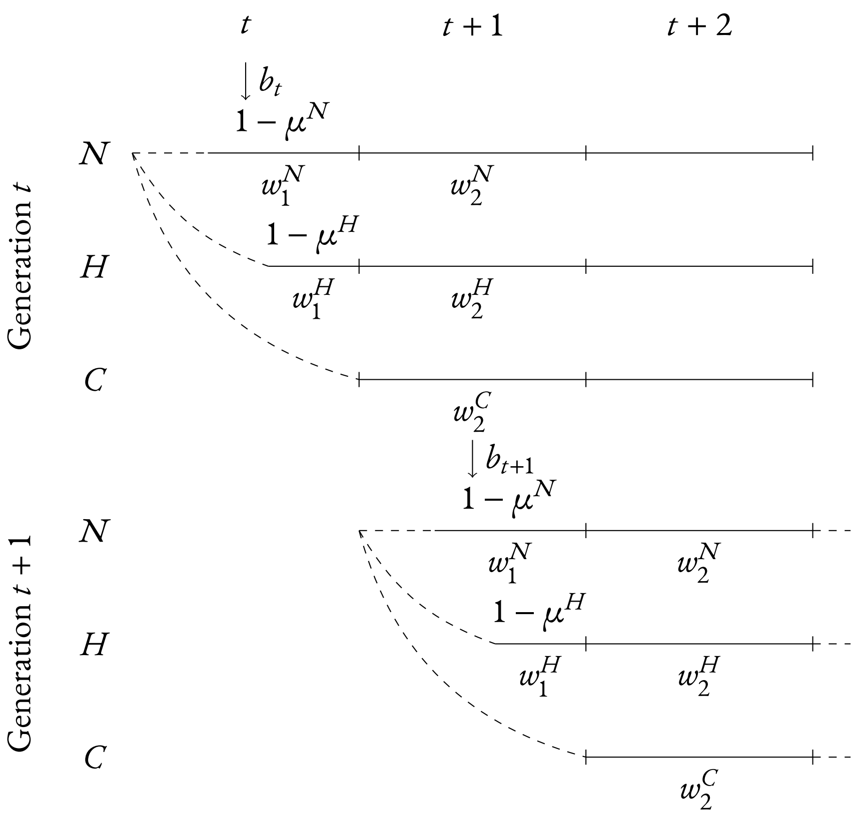
$$y_{1,t} = \begin{cases} (1-\mu_N)w_{N,1} & \text{Dropout} \\ (1-\mu_H)w_{H,1} & \text{High School} \\ 0 & \text{College} \end{cases}$$

where $w_{N,1} < w_{H,1}$ and $0 < \mu_N < \mu_H < 1$, and in the second period,

$$y_{2,t+1} = \begin{cases} w_{N,2} & \text{Dropout} \\ w_{H,2} & \text{High School} \\ w_{C,2} & \text{College} \end{cases}$$

where $w_{N,2} < w_{H,2} < w_{C,2}$. Here μ_N denotes the time devoted to education in the first period of one's life if one drops out of high school, and μ_H the time devoted to education if one graduates from high school. $w_{X,i}$ is the wage level of

Figure 2.3.: *Generation Overlap*



NOTE: Generation overlap and incomes by education level in the model.

an agent with an education level X in period i of his life. I assume $w_{N,1} < w_{H,1}$ and $w_{N,2} < w_{H,2} < w_{C,2}$ in order to have productivity increase with education. $\mu_N > \mu_H$ implies that by dropping out of high school, a young person has more time to work during their youth. I assume

$$(1 - \mu_N) w_{N,1} > (1 - \mu_H) w_{H,1} \tag{2.1}$$

in order to incentivise drop out of high school.

The utility of the agent is of the following form:

$$U = \log(c_{1,t}) + \beta \log(c_{2,t+1}) + \beta^2 \log(c_{3,t+2}) + \gamma\beta \log(b_{t+1}), \tag{2.2}$$

where $\beta \in (0, 1)$ is the subjective discount rate and $c_{1,t}$, $c_{2,t+1}$, and $c_{3,t+2}$ are, respectively, the consumption levels in the first, second, and third period of life. The parents are altruistic towards their children in a joy of giving way and γ captures the degree of altruism towards the child. b_{t+1} is a transfer from the parent to the child in the first period of the child's life.

The agent can save in order to transfer income to later periods. $s_{1,t}$ and $s_{2,t+1}$ denote the savings of young and adults, respectively, to the next period of their life and $r \in (0, 1)$ is the corresponding interest rate on saving, with $R = 1 + r$. In the first period, the agent receives the transfer b_t from his parents and earns net income $(1 - \tau) y_{1,t}$, where τ is the income tax rate. If he acquires an education level X, he has to pay the cost e^X of his education, which is subsidised by the state in the height of h^X. In order to simplify notation, I define $g_1^X = e^X - h^X$. He uses the rest of his income for consumption $c_{1,t}$ and savings $s_{1,t}$. In the second period, he gets the return from first period's savings and net-income $(1 - \tau) y_{2,t+1} - g_2^X$, where g_2^X is an additional tax to induce progressive taxation. g_1^X and g_2^X depend on the education level X in order to introduce progressive taxation and education specific costs and subsidies. The agent uses this for consumption $c_{2,t+1}$, savings $s_{2,t+1}$ and the transfer b_{t+1} to his child. The retired agent consumes all of his savings in the third period. Hence, one gets the following budget constraints:

$$c_{1,t} + s_{1,t} = b_t + (1 - \tau) y_{1,t} - g_1^X, \tag{2.3}$$

$$c_{2,t+1} + b_{t+1} + s_{2,t+1} = (1 - \tau) y_{2,t+1} - g_2^X + R s_{1,t}, \tag{2.4}$$

and

$$c_{3,t+2} = Rs_{2,t+1}. \tag{2.5}$$

The general idea is that agents face a trade off between income in the first period of their lives and higher income in the second period of their lives due to higher education. Agents face a credit constraint $s_{1,t} \geq 0$, thus if their parents do not provide them with a large enough transfer b_t, they cannot optimise their utility over both periods. Thus, agents that receive a small transfer from their parents face a trade off between higher first period consumption and low education, and lower first period consumption and higher education. In order to induce this trade off I assume that life-time income is increasing with education:

$$\begin{aligned}(1-\tau)(Ry^N_{1,t} + y^N_{2,t+1}) - Rg^N_1 - g^N_2 < (1-\tau)(Ry^H_{1,t} + y^H_{2,t+1}) \\ - Rg^H_1 - g^H_2,\end{aligned} \tag{2.6}$$

and

$$(1-\tau)(Ry^H_{1,t} + y^H_{2,t+1}) - Rg^H_1 - g^H_2 < (1-\tau)y^C_{2,t+1} - Rg^C_1 - g^C_2. \tag{2.7}$$

The problem of the agent is to maximise utility (2.2) subject to (2.3) – (2.5). From the first order condition, I obtain the optimal consumption, transfer, and savings for unconstrained agents with $s_{1,t} \geq 0$ and education X:

$$\begin{aligned}
c^{X*}_{1,t} &= \frac{1}{R(1+\beta+\beta^2+\beta\gamma)} \left[(1-\tau)(Ry_{1,t} + y_{2,t+1}) - g^X_2 \right. \\
&\quad \left. + R(b_t - g^X_1)\right], && (2.8) \\
c^{X*}_{2,t+1} &= \frac{\beta}{1+\beta+\beta^2+\beta\gamma} \left[(1-\tau)(Ry_{1,t} + y_{2,t+1}) - g^X_2 \right. \\
&\quad \left. + R(b_t - g^X_1)\right], && (2.9) \\
c^{X*}_{3,t+2} &= \frac{R\beta^2}{1+\beta+\beta^2+\beta\gamma} \left[(1-\tau)(Ry_{1,t} + y_{2,t+1}) - g^X_2 \right. \\
&\quad \left. + R(b_t - g^X_1)\right], && (2.10) \\
& && (2.11)
\end{aligned}$$

$$b_{t+1}^{X*} = \frac{\beta\gamma}{1+\beta+\beta^2+\beta\gamma}\left[(1-\tau)(Ry_{1,t}+y_{2,t+1}) - g_2^X \right.$$
$$\left. +R(b_t - g_1^X)\right], \tag{2.12}$$

and

$$s_{1,t}^{X*} = (1-\tau)y_{1,t} - g_1^X + b_t - c_{1,t}^{X*}. \tag{2.13}$$

As the model features homothetic preferences, in absence of the credit constraint the maximisation of utility is equivalent to the maximisation of life-time income, which can be seen in (2.8) – (2.12). Therefore, if he is not credit constrained, the agent always prefers to have a college education. This follows from (2.6) – (2.7). He only drops out of education because he faces a credit constraint $s_{1,t} \geq 0$, which depends on b_t, as follows from (2.13). Because the income depends on the education decision, I obtain three different thresholds for the transfer $\{\hat{b}_t^N, \hat{b}_t^H, \hat{b}_t^C\}$, below which the agent is facing a binding credit constraint:

$$\begin{aligned}
\hat{b}_t^N &= \frac{1}{R(\beta+\beta^2+\beta\gamma)}\left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - Rg_1^N - g_2^N\right] \\
&\quad - \frac{(1+\beta+\beta^2+\beta\gamma)}{\beta+\beta^2+\beta\gamma}\left[(1-\tau)y_{1,t}^N - g_1^N\right], \\
\hat{b}_t^H &= \frac{1}{R(\beta+\beta^2+\beta\gamma)}\left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - Rg_1^H - g_2^H\right] \\
&\quad - \frac{(1+\beta+\beta^2+\beta\gamma)}{\beta+\beta^2+\beta\gamma}\left[(1-\tau)y_{1,t}^H - g_1^H\right],
\end{aligned} \tag{2.14}$$

and

$$\begin{aligned}
\hat{b}_t^C &= \frac{1}{R(\beta+\beta^2+\beta\gamma)}\left[(1-\tau)y_{2,t+1}^C - Rg_1C - g_2^C\right] \\
&\quad - \frac{(1+\beta+\beta^2+\beta\gamma)}{\beta+\beta^2+\beta\gamma}g_1^C.
\end{aligned} \tag{2.15}$$

It follows from (2.1), (2.6), and (2.7) that $\hat{b}_t^N < \hat{b}_t^H < \hat{b}_t^C$. Thus the higher the education level, the higher the transfer needed in order to not face a binding credit constraint.

The consumption levels of a credit constrained agent are:

$$\bar{c}^X_{1,t} = (1-\tau)y_{1,t} - g^X_1 + b_t,$$

$$\bar{c}^X_{2,t+1} = \frac{\beta}{\beta+\beta^2+\beta\gamma}\left[(1-\tau)y_{2,t+1} + g^X_2\right], \tag{2.16}$$

$$\bar{c}^X_{3,t+2} = \frac{R\beta^2}{\beta+\beta^2+\beta\gamma}\left[(1-\tau)y_{2,t+1} - g^X_2\right], \tag{2.17}$$

and

$$\bar{b}^X_{t+1} = \frac{\beta\gamma}{\beta+\beta^2+\beta\gamma}\left[(1-\tau)y_{2,t+1} - g^X_2\right]. \tag{2.18}$$

Thus an agent who receives a transfer below $\hat{b}^N_t$ will always face a binding credit constraint, irrespective of education, an agent who receives a transfer between $\hat{b}^N_t$ and $\hat{b}^H_t$ can either drop out of high school and be not bounded by the credit constraint or graduate from high school and face a binding credit constraint, an agent who receives a transfer between $\hat{b}^H_t$ and $\hat{b}^C_t$ can either work after high school and not be bounded by the credit constraint or graduate from college and face a binding credit constraint, and an agent receiving a transfer above $\hat{b}^C_t$ can graduate from college without facing a binding credit constraint.

2.3. Education Decision

Agents are identical in this model except for their parental background, i. e. the transfer b_t which they receive from their parents. Thus, their decision on which level of education to attain will only depend on b_t. If they attain higher education, they will have a higher second period income but a lower first period income. This matters only if the agents are credit constrained. I will first consider the decision between high school graduation and drop out of high school and later the decision between high school graduation and college education.

In order to determine whether it is optimal for a young agent to graduate or to drop out of high school, I consider the value of these two option in terms of utility. I define $V^N(b_t)$ as the utility when dropping out and $V^H(b_t)$ as the utility when graduating from high school. When the agent receives a $b_t \in [\hat{b}^N, \hat{b}^H)$ he would face a binding credit constraint when graduating from high school, but not when dropping out. The value of dropping out of high school in terms of utility is

$$V^N(b_t) = \log c_{1,t}^{N*} + \beta \log c_{2,t+1}^{N*} + \beta^2 \log c_{3,t+2}^{N*} + \beta\gamma \log b_{t+1}^{N*},$$

with $c_{1,t}^{N*}$, $c_{2,t+1}^{N*}$, $c_{3,t+2}^{N*}$, and b_{t+1}^{N*} being the consumption and transfer obtained when he dropped out from school. This can be rewritten as

$$\begin{aligned} V^N(b_t) = {} & (1 + \beta + \beta^2 + \beta\gamma) \log \left(\frac{1}{1 + \beta + \beta^2 + \beta\gamma} \right) \\ & + \left[\log \left(\frac{1}{1+r} \right) + \beta \log(\beta) + \beta^2 \log(R\beta^2) + \beta\gamma \log(\beta\gamma) \right] \\ & + (1 + \beta + \beta^2 + \beta\gamma) \log \left[(1-\tau)(R y_{1,t}^N + y_{2,t+1}^N) \right. \\ & \left. - g_2^N + R(b_t - g_1^N) \right]. \end{aligned}$$

The value of graduating from high school in terms of utility is equal to

$$\bar{V}^H(b_t) = \log \bar{c}_{1,t}^H + \beta \log \bar{c}_{2,t+1}^H + \beta^2 \log \bar{c}_{3,t+2}^H + \beta\gamma \log \bar{b}_{t+1}^H,$$

with $\bar{c}_{1,t}^H$, $\bar{c}_{2,t+1}^H$, $\bar{c}_{3,t+2}^H$ and $\bar{b}_{t+1}^H$ being the consumption and transfer of a credit constraint agent with high school income. This can be rewritten to

$$\begin{aligned} \bar{V}^H(b_t) = {} & (\beta + \beta^2 + \beta\gamma) \log \left(\frac{1}{\beta + \beta^2 + \beta\gamma} \right) \\ & + \left\{ \log \left[(1-\tau) y_{1,t}^H - g_1^H + b_t \right] + \beta \log \beta + \beta^2 \log(R\beta^2) + \beta\gamma \log(\beta\gamma) \right\} \\ & + (\beta + \beta^2 + \beta\gamma) \log \left[(1-\tau) y_{2,t+1}^H - g_1^H \right]. \end{aligned}$$

I am looking for the value of the transfer $b_t^H \in [\hat{b}_t^N, \hat{b}_t^H)$ above which the agent finishes high school and below which he drops out of high school. I obtain this cutoff value by looking at the b_t for which the agent is indifferent between graduating from high school and dropping out, i. e. for which the values of the two value functions are equal. I get for the transfer b_t^H

$$V^N(b_t^H) = \bar{V}^H(b_t^H).$$

This can be reformulated as

$$
\begin{aligned}
&\log \left\{ \frac{(1-\tau)(Ry^N_{1,t} + y^N_{2,t+1}) - g^N_2 + R(b^H_t - g^N_1)}{R(1+\beta+\beta^2+\beta\gamma)\left[(1-\tau)y^H_{1,t} - g^H_1 + b^H_t\right]} \right\} \\
&\quad = (\beta+\beta^2+\beta\gamma)\log\left[\frac{1+\beta+\beta^2+\beta\gamma}{\beta+\beta^2+\beta\gamma}\right. \\
&\qquad \left. \cdot \frac{(1-\tau)y^H_{2,t+1} - g^H_2}{(1-\tau)(Ry^N_{1,t} + y^N_{2,t+1}) - g^N_2 + R(b^H_t - g^N_1)}\right],
\end{aligned} \tag{2.19}
$$

where b^H_t is the transfer that fulfils the equation above. Agents receiving a transfer above b^H_t graduate from high school, agents receiving a transfer below b^H_t drop out of high school.

I next proceed in a similar way to obtain the threshold b^C_t above which agents graduate from college. In particular, I compare the value of high school education in terms of utility for an agent not facing a binding credit constraint

$$
\begin{aligned}
V^H(b_t) &= (1+\beta+\beta^2+\beta\gamma)\log\left(\frac{1}{1+\beta+\beta^2+\beta\gamma}\right) \\
&\quad + \left[\log\left(\frac{1}{1+r}\right) + \beta\log(\beta) + \beta^2\log(R\beta^2) + \beta\gamma\log(\beta\gamma)\right] \\
&\quad + (1+\beta+\beta^2+\beta\gamma)\log\left[(1-\tau)(Ry^H_{1,t} + y^H_{2,t+1})\right. \\
&\quad \left. - g^H_2 + R(b_t - g^H_1)\right],
\end{aligned}
$$

with the value of college education for an agent facing a binding credit constraint

$$
\begin{aligned}
\bar{V}^C(b_t) &= (\beta+\beta^2+\beta\gamma)\log\left(\frac{1}{\beta+\beta^2+\beta\gamma}\right) \\
&\quad + \left[\log(b_t - g^C_1) + \beta\log\beta + \beta^2\log(R\beta^2) + \beta\gamma\log(\beta\gamma)\right] \\
&\quad + (\beta+\beta^2+\beta\gamma)\log\left[(1-\tau)y^C_{2,t+1} - g^C_2\right].
\end{aligned}
$$

Receiving $b^C_t \in [\hat{b}^H_t, \hat{b}^C_t)$ makes an agent indifferent between high school and college education if it fulfils the following condition:

$$
V^H(b^C_t) = \bar{V}^C(b^C_t),
$$

which can be reformulated as

$$\log\left[\frac{(1-\tau)(Ry_{1,t}^{H}+y_{2,t+1}^{H})-g_{2}^{H}+R(b_{t}^{C}-g_{1}^{H})}{R(1+\beta+\beta^{2}+\beta\gamma)(b_{t}^{C}-g_{1}^{C})}\right]$$
$$=(\beta+\beta^{2}+\beta\gamma)\log\left[\frac{1+\beta+\beta^{2}+\beta\gamma}{\beta+\beta^{2}+\beta\gamma}\right.$$
$$\left.\cdot\frac{(1-\tau)y_{2,t+1}^{C}-g_{2}^{C}}{(1-\tau)(Ry_{1,t}^{H}+y_{2,t+1}^{H})-g_{1}^{H}+R(b_{t}^{C}-g_{1}^{H})}\right]. \tag{2.20}$$

An agent that receives a b_t larger than b_t^C will go to college, an agent that receives a b_t smaller than b_t^C will not go college and start working after graduating from high school.

Thus depending on the value of b_t, the agent can be in six different situations: (i) if $b_t < \hat{b}_t^N$ he drops out of high school but is still facing a binding credit constraint, (ii) if $b_t \in [\hat{b}_t^N, b_t^H)$ he drops out of high school but is not facing a binding credit constraint, (iii) if $b_t \in [b_t^H, \hat{b}_t^H)$ he graduates from high school but is facing a binding credit constraint in doing so, (iv) if $b_t \in [\hat{b}_t^H, b_t^C)$ he graduates from high school and is not facing a binding credit constraint in doing so, (v) if $b_t \in [b_t^C, \hat{b}_t^C)$ he goes to college and faces a binding credit constraint in doing so, and (vi) if $b_t \geq \hat{b}_t^C$ he graduates from college without facing a binding credit constraint.

2.4. Wage Polarisation and Social Mobility

In order to understand how trends in wages and policies affect intergenerational mobility, I use (2.19) and (2.20) to perform a comparative statics exercise on the transfers needed to graduate from high school and college, b^H and b^C, respectively. If these minimum transfers increase, ceteris paribus, less student will receive a transfer above them and there will be less upward social mobility. If these minimum transfers decrease, more students will receive a transfer above these minimum transfers and there will be more upward social mobility. One can use that to make predictions on how changes in wage differentials affect the graduation rates of high school and college, which I will use in Section 2.6 to propose policies oriented at increasing graduation rates.

The choice of whether to acquire higher education or not is defined by the opportunity cost of education and the return on education. Because $b^H < b^C$, and anyone who wants to go to college needs a high school degree, I only compare the incomes of the "neighbouring" education levels, i. e. the opportunity costs and return on education of high school graduates relative to drop outs, and the opportunity cost and return on education of high school graduates and college graduates.

I define the opportunity cost of education as how much lower first period income is when spending more time on education. In the framework of this model, I define therefore the opportunity cost of high school education as how much higher the first period income is when dropping out than when graduating from high school. When an agent drops out of high school, he earns $y^N_{1,t}$, so the opportunity cost is this income relative to the income he would earn if he would graduate from high school, i. e. $y^N_{1,t}/y^H_{1,t}$. Since college graduates cannot work in the first period of their lives in this model, the opportunity cost of college education is the income the student could earn in the first period of his life when not going to college, i. e. the income as a young high school graduate $y^H_{1,t}$. I define the returns on education as how much higher second period income is when graduating. In the framework of this model, the return on high school education is how much relatively higher second period high school graduate income is than drop out income. If the agent drops out of high school in the first period of his life, he earns $y^N_{2,t+1}$ in the second period of his life. If he graduates from high school, he earns $y^H_{2,t+1}$. Thus the return on high school education is $y^H_{2,t+1}/y^N_{2,t+1}$. The return on college education is in this model how much higher second period income of college graduates is relative to the income of high school graduates. If the agent only graduates from high school and does not go to college in the first period of his life, he earns $y^H_{2,t+1}$ in the second period of his life. If he graduates from college instead, he earns $y^C_{2,t+1}$ in the second period of his life. Thus the return on college education is $y^C_{2,t+1}/y^H_{2,t+1}$.

In order to analyse the role of these wage premiums, I establish the following relationship between $c^{N*}_{1,t}$, $\bar{c}^H_{1,t}$, $c^{H*}_{1,t}$ and $\bar{c}^C_{1,t}$:

Lemma 1. *If an agent receives b^H_t then $c^{N*}_{1,t} > \bar{c}^H_{1,t}$, and if an agent receives b^C_t then $c^{H*}_{1,t} > \bar{c}^C_{1,t}$.*

Proof. Assume that if $b_t = b^H_t$ then $c^{N*}_{1,t} \leq \bar{c}^H_{1,t}$. It follows from (2.6), (2.7), (2.9) to (2.12), and (2.16) to (2.18), that $\bar{c}^H_{2,t+1} > c^{H*}_{2,t+1} > c^{C*}_{2,t+1}$, $\bar{c}^H_{3,t+2} > c^{H*}_{3,t+2} > c^{C*}_{3,t+2}$, and $\bar{b}^H_{t+1} > b^{H*}_{t+1} > b^{C*}_{t+1}$. This implies that $V^N(b_t) < \bar{V}^H(b_t)$. This is in contradiction to the definition of b^H_t, thus $c^{N*}_{1,t}$ has to be larger than $\bar{c}^H_{1,t}$. Using the same logic, one can derive that $c^{H*}_{1,t} > \bar{c}^C_{1,t}$. □

First, I use the implicit function theorem to get the effect of a change in the opportunity cost of high school education on the value of b_t^H. In order to simplify notation, I use in the following

$$\Gamma = \beta + \beta^2 + \beta\gamma.$$

Proposition 1. *An increase in the opportunity cost of high school education $y_{1,t}^N / y_{1,t}^H$ increases the transfer b_t^H needed in order to graduate from high school.*

Proof. The partial derivative of b_t^H with respect to $y_{1,t}^N / y_{1,t}^H$ is

$$\frac{\partial b_t^H}{\partial \frac{y_{1,t}^N}{y_{1,t}^H}} = -(1+\Gamma)(1-\tau)y_{1,t}^H \left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]$$
$$\cdot \left\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right] - \left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\right\}^{-1}.$$

It follows from Lemma 1 that this is positive. □

Thus an increase in the opportunity cost of high school education increases the transfer needed for graduating from high school and ceteris paribus less students will graduate. If $y_{1,t}^N / y_{1,t}^H$ increases, the relative value of the outside option to graduating from high school increases, which makes it more costly to graduate from high school for those agents facing a binding credit constraint. Thus, the transfer that these agents need to get in order to be indifferent between graduating from high school and dropping out has to be higher. The effect of the opportunity costs on high school education on b_t^H is the larger, the higher first period income for dropouts $y_{1,t}^N$, the higher life-time income for dropouts.

The transfer needed to graduate from high school is affected by the wage premium on high school graduation in the following way:

Proposition 2. *An increase in the return on high school education $y_{2,t+1}^H / y_{2,t+1}^N$ decreases the transfer needed in order to graduate from high school b_t^H.*

Proof. The partial derivative of b_t^H with respect to $y_{2,t+1}^H / y_{2,t+1}^N$ is

$$\begin{aligned}\frac{\partial b_t^H}{\partial \frac{y_{2,t+1}^H}{y_{2,t+1}^N}} = {} & \Gamma(y_{2,t+1}^N - g_2^N)\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right] \\ & \cdot \left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right] \\ & \cdot (y_{2,t+1}^H - g_2^H)^{-1}\left\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\right. \\ & \left. - \left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\right\}^{-1}.\end{aligned}$$

which following Lemma 1 is negative. □

Thus, an increase in the return on high school education will lead to a decrease in the transfer needed in order to graduate from high school. If the returns on high school education increase, it becomes more attractive to graduate relative to the option of dropping out of high school. The transfer needed to be indifferent between dropping out and graduating from high school will be lower and, ceteris paribus, more students will receive a transfer above this threshold, leading to an increase in the number of students graduating from high school. An increase in the return on graduating from high school decreases the transfer needed in order to graduate from high school, and thus ceteris paribus increases the number of high school graduates.

The effects of changes in the wage premiums for college education can be analysed in the same way. First I look at the opportunity cost of education. As college graduates cannot work in the first period of their lives, the opportunity cost of college education is equal to $y_{1,t}^H$.

Proposition 3. *An increase in the opportunity costs of college education $y_{1,t}^H$ increases the transfer b_t^C needed in order to graduate from college.*

Proof. The partial derivative of b_t^C with respect to $y_{1,t}^H$ is equal to

$$\frac{\partial b_t^C}{\partial y_{1,t}^H} = -\frac{(1-\tau)Rb_t^C}{(1+\Gamma)R(b_t^C - g_1^C) - \left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_1^C + R(b_t^C - g_2^C)\right]},$$

which following Lemma 1 is positive. □

An increase in the wage of young high school graduates increases the value of the alternative to graduating from college. Thus, going to college becomes relatively less attractive and the transfer needed to be indifferent between college and only high school education increases. This leads to a decrease in the number of agents receiving a transfer above this threshold. It has no effect on the number of college graduates that are not facing a binding credit constraint, thus an increase in $y_{1,t}^H$ will only decrease the number of college graduates facing a binding credit constraint.

The return on college education has the following effect on the transfer b_t^C needed for graduating from college:

Proposition 4. *An increase in the return on college education $y_{2,t+1}^C / y_{2,t+1}^H$ decreases the transfer b_t^C needed in order to graduate from college.*

Proof. The partial derivative of b_t^C with respect to $y_{2,t+1}^C / y_{2,t+1}^H$ is equal to

$$\frac{\partial b_t^C}{\partial \frac{y_{2,t}^C}{y_{2,t}^H}} = \frac{\Gamma(y_{2,t+1}^H - g_2^H)(b_t^C - g_1^C)}{y_{2,t+1}^C - g_2^C} \cdot \frac{(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_2^H + R(b_t^C - g_1^H)}{(1+\Gamma)R(b_t^C - g_1^C) - \left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_1^H + R(b_t^C - g_2^H\right]},$$

which following Lemma 1 is negative. □

Thus an increase in the return on college education decreases the transfer needed in order to graduate from college. It increases the value of having college education relative to the value of only high school education, making college education more attractive and thus the transfer needed in order to be indifferent between the two is lower. Hence there will be more agents receiving a transfer above this and more graduating from college. This is the larger the larger $y_{2,t+1}^H$ and b_t^C and the first period income of a high school graduate are. It is the smaller the larger $y_{2,t}^C$.

To summarise the results of the comparative static analysis: A decrease in the opportunity cost of education increases the number of graduates of a higher education degree. An increase in the return on education increases the number of graduates. A detailed analysis of the components can be found in Appendix A.1. If a decrease in the opportunity cost of education is due to a decrease in the wage of young dropouts, the increase in graduates will be entirely due to an increase of graduates facing a binding credit constraint, with the number of agents graduating that are not facing a binding credit constraint staying the same. This is due to the fact that a change in the wage of dropouts affects the thresholds $\{b_t^H, b_t^C\}$ above which agents graduate, but does not enter in the definitions (2.14) and (2.15) of the thresholds $\{\hat{b}_t^H, \hat{b}_t^C\}$ that govern whether or not the budget constraint of an agent is binding. If a decrease in the opportunity cost of education is due to an increase in the wage of young graduates, the total number graduates and the number of graduates not facing a binding credit constraint increases. The wage of young graduates also enters into (2.14) and (2.15), and thus $\{\hat{b}_t^H, \hat{b}_t^C\}$ are affected by a change of them. For the same reasons, an increase in the return on education that is due to a decrease in the wage of old drop outs increases the total number of graduates, but does not affect the number of graduates not facing a binding credit constraint. If an increase in the return on education are due to an increase in the wage of graduates, then the total number of graduates increases and the number of graduates facing a binding credit constraint increases, while the number of graduates not facing a binding credit constraint decreases.

Labour market polarisation means that the wages of middle skilled workers decrease and the wages of low and high skilled workers increase. As shown above, a decrease in the high school premium implies less mobility from the lower income level as b^H increases. It also implies an increase in the college premium, which implies that b^C decreases. Based on this comparative statics I expect to see a negative relationship between the opportunity cost of education and graduation rates, and a positive relationship between the returns on education and graduation rates. In the next Section, I will test these predictions on U. S. high school dropout rates.

2.5. Empirical Analysis

The model developed in this paper provides an interpretation of wage premiums as opportunity costs of education and returns on education. The comparative static analysis shows that an increase in the return on education increases educational attainment, whereas an increase in the opportunity costs of education decreases educational attainment. Any policy that is aimed at changing one wage differential must be financed and thus might have consequences for the other wage differential. Thus in order to make policy recommendations, I estimate their relative importance.

I focus on the drop out rate from high school, because as discussed in Section 2.1 high school education is important to explain differences in upward social mobility between countries, and as I expect mobility between regions for educational purposes to be less of an issue for high school education than for college education. I regress the drop out rate from high school on the wage premiums defined in the previous section in order to estimate the relative importance of these wage premiums. Analogous to Chetty et al. (2014a), I use commuting zones as defined by Tolbert and Sizer (1996) as unit of observations. I use the crosswalk files provided by Autor and Dorn (2013) to aggregate the average wages by education level, gender, and age as well as the high school dropout rate at commuting zone level.

For the wage data, I use the 1990 U. S. 5% census IPUMS. I define agents as "young" if they are of the age of 22 or below (graduating from high school at the age of 18 plus 4 year of college education). For the dropout of high school, the data come from the common core of data from the IES NCES in the year 2001.

I cannot directly observe the life period incomes $y_{1,t}$ or $y_{2,t+1}$ as defined in the model, nor the time spent in education μ, but I can observe the individual wage income in one year, which in the model is positively linearly related to the live period incomes. I use the average wage income by age group in one region in the census year w_1 and w_2 as a measure of life period income. As in the model, I define the opportunity cost of education as w_1^N/w_1^H, and the return on education as w_2^H/w_2^N. I regress these wage premiums by gender on the dropout rate $P(D)$ by gender, controlling for the logarithm of the average household income (I use the one derived by Chetty et al. (2014a)) in the commuting zone $\log(I_{HH})$ as a measure of overall wealth in the region.

The result of this regression is presented in Table 2.5. I find a significant effect of both the opportunity cost of education and the return on education on the dropout rate for men with the signs predicted by the model. A 10 percentage point increase in the wage differential between high school dropouts and high school graduates of young men (the opportunity cost of high school education) increases the high school dropout rate by 0.7 percentage points. A 10 percentage point increase in the wage differential between high school graduates and high school dropouts of old men (the return on education) decreases the high school dropout rate by 0.2 percentage points. This implies that changes in the relative wage of young graduates have a more than three times higher impact on the high school dropout rate.

However, I do not find any significant effect of wage premiums on the dropout rate for women. One reason that there is no effect of the wage premiums for women might be the importance of childbirth and its effect of wages and labour force participation.

Table 2.5.: *High School Dropout by Gender*

	(1) Men	(2) Women
w_1^N/w_1^H	0.0687***	-0.00154
	(0.0221)	(0.0170)
w_2^H/w_2^N	-0.0192***	-0.00362
	(0.00600)	(0.00296)
$\log(I_{HH})$	-0.0434***	-0.0231***
	(0.0105)	(0.00868)
Constant	0.527***	0.293***
	(0.114)	(0.0941)
Observations	207	207
R^2	0.155	0.034
Adjusted R^2	0.143	0.020

NOTE: Std. deviations in parenthesis.
$^{*}p < 0.10$ $^{**}p < 0.05$
$^{***}p < 0.01$

2.6. Policy Analysis

I now use the model and the empirical results to draw some conclusions in order to propose policies that increase educational attainment of children from low income families. The analysis of the comparative statics of the model introduced in this paper suggests the following: One can increase the number of children attaining educational degrees by either decreasing the opportunity costs or by increasing the return on education. Both are possible by two ways, by either affecting the wages of those who drop out of education or by affecting the wages of those who attain education. A government could introduce policies that aim at decreasing the after tax income of dropouts $\{(1-\tau)y_{1,t}^N - g_1^N; (1-\tau)y_{2,t+1}^N - g_2^N\}$ by taxing these incomes higher (e. g. by introducing less progressive taxation schemes) and thus decreases the opportunity cost to education and increase the return on education. This would decrease the transfer needed in order to graduate from high school, but it would also decrease the income of poor parents and thus decrease the transfer received by poor children. It would also imply a redistribution of income from poor households to richer households.

Based on the empirical assessment, in which the opportunity cost to high school education had a much higher impact on high school dropout rates, I propose a different policy: Increasing the after tax income of young high school graduates $(1-\tau)y_{1,t}^{H}-g_{1}^{H}$ financed by decreasing the income of older high school graduates $(1-\tau)y_{2,t+1}^{H}-g_{2}^{H}$. This means that the opportunity cost of education decrease but also the return on education decreases. Note that in practice this implies a transfer from relatively rich adult individuals to poor young individuals

This would not only increase the number of high school graduates but also increase the number of high school graduates that are not facing a binding credit constraint. As the government cannot directly affect the market wages w_{1}^{H} and w_{2}^{H}, such a policy could be achieved by increasing the subsidy for high school education h, resulting in a decrease in g_{1}^{H}, and increase the tax of high school graduates g_{2}^{H}.

Proposition 5. *An increase in the subsidy for high school education h^{H} (a decrease in g_{1}^{H}) payed by an increase in the taxes for old high school graduates g_{2}^{H} of equal size leads to a decrease in the transfer needed in order to graduate from high school and a decrease in the transfer needed in order to be not facing a binding credit constraint when graduating from high school $\hat{b}_{t}^{H}$.*

Proof. See Appendix A.2.

I find in the empirical analysis that the wage differential for young high school graduates has a more than three times higher impact on the high school dropout rate. At the same time, the wage level of young people is much lower than the one of old people, and also the period working being young is shorter, thus just a small percentage decrease in the return on education can cause a larger percentage decrease in the opportunity cost of education, leading to an overall decrease in the dropout rate.

I do not have data on the importance of wage differentials of college graduates, but the analysis of our model points to a similar policy of decreasing the return on education in favour of decreasing the opportunity cost of education in order to increase college graduation rates and decrease the number of college graduates not facing a binding credit constraint. This is especially relevant in the debate on tuition fees. Proponents of tuition fees for tertiary education argue that such tuition fees induce a payment by the beneficiary. This model suggests, that it might be preferable to pay for tertiary education after graduation, i. e. through higher taxes or publicly provided and subsidised student loans, and to subsidise tertiary education for poor students in order to decrease the opportunity costs of education.

This might not only increase graduation rates of poor students, but also decrease the high levels of student debt in the U. S. and U. K.

Again, the government cannot directly change w_2^C, but it can set the education subsidy h^C and the tax rate for college graduates g_2^C. The suggested policy would be to finance an increase in h^C (a decrease in g_1^C) through an increase in g_2^C.

Proposition 6. *An increase in the subsidy for college education h^C (a decrease in g_1^C) payed by an increase in the taxes for old college graduates g_2^C leads to a decrease in the transfer needed in order to graduate from college and a decrease in the transfer needed in order to be not facing a binding credit constraint when graduating from college $\hat{b}_t^C$.*

Proof. See Appendix A.2.

Such a policy would increase the number of college graduates while decreasing the number of college graduates facing a binding credit constraint.

2.7. Conclusion

This paper investigates the effect of wage differentials on intergenerational income mobility by introducing a Galor and Zeira (1993) type OLG model with three levels of education where children make their own educational choices base on a transfer they receive from their parents. In this model, there are two forces that define the educational choices: the return on education and the opportunity cost of education.

In an empirical assessment of these two forces, I find that the opportunity cost of education has a much stronger influence on the probability to finish high school for men than the return on education. I use the conclusions drawn from the model to propose a policy which reduces the opportunity cost of education by subsidising education payed by taxes on older graduates. This policy has the advantage that it increases educational attainment of the poor, increases the number of graduates not facing a binding credit constraint, decreases income inequality due to age differences, while not implying changes in life-time income nor redistribution between education groups.

This paper argues that education and distributional policies should not only be concerned about their effect on the return on education, but also about their effect on the opportunity cost of education for the students. Policies that carefully manage age and educational premiums can improve educational outcomes for poor children, increasing upward mobility and intergenerational social mobility in general. It would be interesting to explore this insights further by modelling the role of ability into the model, which would allow to identify differences in the ability distribution of children with the same education level from different parental backgrounds. It would also be interesting to investigate gender differences in educational attainment. I do not find any significant effect of wage differentials on the high school dropout rate of women. It would be interesting to explore whether a similar model including endogenous fertility might provide an explanation of educational patterns of women.

3 Explaining Trends in Fertility and Childlessness in Germany

3.1. Introduction

During the 20th century, Germany experienced a strong decline in fertility. Between the cohorts of women born in the 1930s and the cohorts born in the early 1960s, the completed fertility rate (the number of children that the average women of a generation has over her lifetime, henceforth CFR) has decreased from 2.2 to 1.6. The literature on demographical change mainly explains this by a Becker (1960) style "quantity/quality" trade-off in the number of children and the education of children. The richer a family becomes, the more important becomes the education that the parents can provide relative to the number of children. This substitution effect between the number of children and their education leads to a decline in fertility with increasing household income. Other explanations like Galor and Weil (1996) and Galor and Moav (2000) highlight the role of supply driven mechanism like the increase in human capital returns or changes in the gender wage gap.

Recently there has been an increasing interest in distinguising between intenive fertility (number of children per women with at least one child) and extensive fertility (childlessness) (see Gobbi, 2013; Baudin et al., 2015, 2018), highlighting the role played by the latter.[1] As I show in Section 3.2 of this paper, the decline in fertility of the cohorts of women born after 1950 is entirely due to an increase in the share of childless women. The quantity/quality framework can produce a decrease in fertility due to a decrease in the number of children per women with at least one child, but it cannot produce childlessness because of the "essential complementarity" between quantity and quality: as Aaronson et al. (2014) show, in order to have a substitution effect of the number of children for investment into their education, there must be at least one child present. This means that in order to have some child "quality", the household has to have at least one child. The

[1] See Baudin et al. (2019) for an introduction to the economic literature on childlessness.

literature on childlessness, on the other hand, has mainly focused on generating childlessness independent of education investment. However, in the literature on the quantity/quality trade-off, this investment into education is the main driver of the mechanism.

This paper aims to reconcile the literature on the quantity/quality trade-off with the literature on childlessness in a single framework. First, I decomposes the decline in fertility in Germany using data from the German Socio Economic Panel (SOEP) and investigates the opportunity cost of childbirth for women. Second, I develop an overlapping generations model that is able to reproduce the decline of both intensive and extensive fertility in order to understand the mechanism behind the decline in fertility. This model features a quantity/quality trade-off and opportunity cost of children in terms of work and agents are heterogeneous with respect to their preference for children. I find in an event study analysis that the main opportunity cost of childbirth is in terms of time, and therefore I feature in the model this particular opportunity costs. I show in a calibration exercise that this model can replicate the decline in the number of children per women with at least one child and the increase in childlessness observed in the data through a decrease in the gender wage gap induced by the increased importance of mental (i. e. non-physical) labour. Therefore I argue that the decrease in the number of children per women and the increase in childlessness is driven by the increase in the relative financial value of working time for women due to a decrease in the gender wage gap.

In order to model the fertility choice of women, I first decompose empirically the decline in fertility and examine the opportunity costs in terms of work associated with children. For this, I use data from the German Socio Economic Panel (SOEP). I decompose the completed fertility rate into two dimensions: extensive fertility (if a woman has children or not) and intensive fertility (how many children a woman with at least one child has). I identify in the empirical analysis two stages of the decline in fertility: until the generations born in the 1950s the fertility declines due to a strong decrease in intensive fertility. The share of women for these birth cohorts that are childless stays more or less constant, but the number of children per woman that has at lest on child is decreasing strongly. For the cohorts born after 1950, the decrease in fertility is entirely due to a decrease in extensive fertility: the number of children per woman that has at least one child is staying constant or even increasing, but the share of women staying childless is increasing. The decline in fertility over the entire period can be explained by 70% by the decrease in extensive fertility, and by around 30% by the decrease in intensive fertility. Therefore I argue that childlessness plays an increasingly important role in understanding the decline in fertility.

Furthermore, I find that women's education expansion plays only a secondary role for the decline in fertility. The fertility trends are very similar for women of different education groups. Even though there was a massive increase in women's education attainment, the increase in women's education can only explain around 25% of the total decline in fertility. The differences in fertility between education groups in the later birth cohorts are entirely due to differences in intensive fertility, with extensive fertility being the same across all groups.

I conclude from this decomposition of the completed fertility rate, that there is a need for a model of fertility and childlessness, as the quantity/quality theory alone cannot explain childlessness. To develop such a model, I empirically investigate the nature of the opportunity costs of children. For this I use the event study approach developed by Kleven et al. (2018) using the SOEP data. As Kleven et al. (2018), I find a substantial decline in labour income of women after the birth of their first child relative to their labour income before childbirth, and no negative change of men's labour income. I show that the relative costs of having children in terms of women's forgone income relative to the income before childbirth is independent of education and driven by a decline in labour market participation and hours worked.[2] There is no significant change in the hourly wage.

Based on these findings, I develop an overlapping generations model with a Galor and Weil (1996) type economy with two types of labour (mental and physical). In this model, women face an opportunity costs of having children in terms of time. Households are heterogeneous in terms of their preference for children. They choose whether or not to have children, and if so the number of children and the investment into their education. In this model, an increase in women's wages (with decreasing gender wage gap) due to an increase in the wages on mental labour (relative to physical labour) leads to an increase in the opportunity cost associated with each child in terms of forgone women's labour income. In a particular household this first leads to a decrease in fertility, and then to childlessness, depending on the strength of the preference for children. Taking the distribution of preferences for children as exogenous, an increase in the wage on mental labour

[2]In a difference in difference estimation comparing women with at least one child to women with no children, I find some evidence that this child penalty is higher relative to the income which a women would have had if she had no children for women with higher education.

decreases aggregate fertility and increases childlessness in the economy. I show in an calibration exercise that this model can produce the overall patterns observed in the fertility decomposition through a decreasing gender wage gap that matches data on labour force participation (though the distribution of preferences for children chosen cannot match the convexity of childlessness and intensive fertility).

This approach contributes to the literature on demographical change and fertility. Understanding the mechanism behind the decline in fertility is of particular importance for developed countries with strongly declining fertility like Germany because welfare state policies interact with fertility, and affect the sustainability of pension systems (Abío et al., 2004). From the point of view of less developed countries this kind of analysis can drive lessons to take advantage of the "demographical dividend" and create welfare state policies that help smooth the demographic transition (Rentería et al., 2016; Sánchez-Romero et al., 2017).[3] [4]

The empirical literature has generally found a negative relationship between a women's education level and the number of children she has. Rindfuss et al. (1980) show that education has a significant influence on age at first birth, but don't find a reverse effect. They argue that almost all of the effect of education on completed fertility goes through age at first birth. Gobbi (2013) shows that both within the U. S. as well as across countries there is a negative relationship between childlessness and completed fertility. Countries with lower levels of fertility also experience higher level of childlessness. Baudin et al. (2015) show that within the U. S. intensive fertility is decreasing with education, whereas extensive fertility features an inverted U-shaped relation to education of women. This means that women with low levels of education have a relatively high probability of being childless, but if they have children they have relatively many. Women with medium levels of education have only a small probability of being childless, but tend to have fewer children if they are mothers than women with lower education. Women with high levels of education are relatively likely to be childless, but also have the lowest intensive fertility. Baudin et al. (2018) show a similar U-shape in childlessness and women's education for developing countries. Baudin et al. (2019) show that a large number of developed countries exhibit a U-shape in childlessness over time, where childlessness is decreasing first and then increasing again.

There is an evolving literature on the opportunity cost of children for women

[3] Sánchez-Romero et al. (2017) and Rentería et al. (2016) investigate the effect of the demographical shift and the education expansion in particular of women to economic development.

[4] See Sardon (2006) for a survey of population trends.

in terms of work, the “child penalty”. Adda et al. (2017) show that women with higher preferences for children choose occupations with lower wages but also lower child penalties. Kleven et al. (2018) show that a decline in income after childbirth is present for women but not for men. It can explain a substantial part of the gender wage gap. Alders and Broer (2005) show in their model that an increase in the opportunity costs of children will lead to a decrease in fertility and an increase in labour supply.

The model developed in this paper is contributing to a large literature on endogenous fertility models: Galor and Weil (2000) explains the demographical change as the outcome of a shift from a Malthusian growth regime with high population growth into a modern growth regime with a negative relationship between the income per capita and population growth. In Galor and Weil (1996), an increase in capital over the developpment of an economy leads to an increase in the wages on mental labour (relative to physical labour) which increases the wages for women relative to men and decreases fertility.

Kremer and Chen (2002) show that countries with higher levels of income inequality have higher fertility differentials between women with high and low education. They argue that higher wages reduce fertility because the substitution effect outweighs the income effect, i. e. that with higher wages women prefer to work instead of having more children. de la Croix and Doepke (2003) argue that because of this, higher income inequality leads to a higher population weight on children from poor backgrounds with low education. Galor and Tsiddon (1997) argue in a similar vein that fertility and the income distribution account for a substantial part of the per-capita output and growth performance differences across countries.

Baudin et al. (2019) argue that there are four types of childlessness: natural sterility, mortality driven childlessness (no surviving children), opportunity driven childlessness driven by the opportunity cost in terms of work for high educated women, and poverty driven childlessness (social sterility). In Baudin et al. (2015), the patterns in observed completed fertility are due to the interaction between the “social sterility” of the poor and the high opportunity costs to children for well educated women. Based on this framework, Baudin et al. (2018) investigates the importance of these causes of childlessness for the completed fertility rate of developing countries.

In Blackburn and Cipriani (2005), the demographical transition is driven by a reversion of the net intergenerational transfers, the so called “Caldwell hypothesis". In low developed countries transfers go from children to the parents, whereas in higher developed it is other way round. This changes the incentives for having children. Galor and Zang (1997) show in an overlapping generations model with exogenous fertility, that higher fertility decreases the per-child resources for investment into their education, decreasing growth rates. This is also explained trough income inequality.

There are a number of papers that explain fertility as the outcome of preferences. Booth and Kee (2006) and Gobbi (2013) model fertility preferences as the outcome of a cultural transmission mechanism in the line of Bisin and Verdier (2001). Both parents and society transmit their preferences for children to their children. In Greenwood et al. (2003), fertility is the outcome of a Nash bargaining between women and men.

The rest of the paper is structured as follows: Section 3.2.1 decomposes the fertility decline in Germany, Section 3.2.2 investigates the opportunity costs of children, Section 3.3 introduces the model, Section 3.4 analyses its explanatory power in a calibration exercise, and Section 3.5 concludes the paper.

3.2. Empirical Analysis

During the 20th century there was a strong decline in the fertility rate in Germany, leading to a drop in childbirth below the replacement rates of around 2.1 that would keep the population stable without immigration. I want to decompose this decline in fertility in order to understand the underlying mechanisms. For this, I use the German socio economic panel (SOEP)[5] which is a panel data set tracking individuals from 1984 to 2015. In order to have data representative of the population, I drop the data from the Familien in Deutschland (FiD) data set that was added to the SOEP in 2015, which would otherwise lead to an oversampling of women who have children. I observe per year on average 17698 individuals, 9842 of which are women, and 354 births. This data allows us to observe a range of indicators for each individual, as well as track them over time to assess the effect of childbirth.

In this section, I first use this data to decompose the completed fertility rate of Germany for the birth cohorts of 1930 to 1965, and then use data on observed

[5] Socio-Economic Panel (SOEP), data for years 1984-2015, version 32, SOEP, 2015, doi:10.5684/soep.v32.

childbirths between 1989 and 2005 to investigate the opportunity cost of childbirth in terms of work.

3.2.1. Decomposing Completed Fertility

Firstly, I look at the completed fertility rate. This is the number of children a women has in her lifetime. It is measured at the age of 45/50, since it is assumed that afterwards women will not give birth anymore. I have data for the birth cohorts from 1930 to 1965. As it can be seen in Figure 3.1, there was a strong decline in the completed fertility rate from 2.2 children per woman for the birth cohort of 1930 to below 1.6 for the birth cohort of 1965. This means a decline in the fertility rate from birth cohorts of the early 1930s to birth cohorts of the early 1960s of about 27%.[6] The trend in the SOEP dataset closely tracks the completed fertility rate from the census (Microzensus).[7]

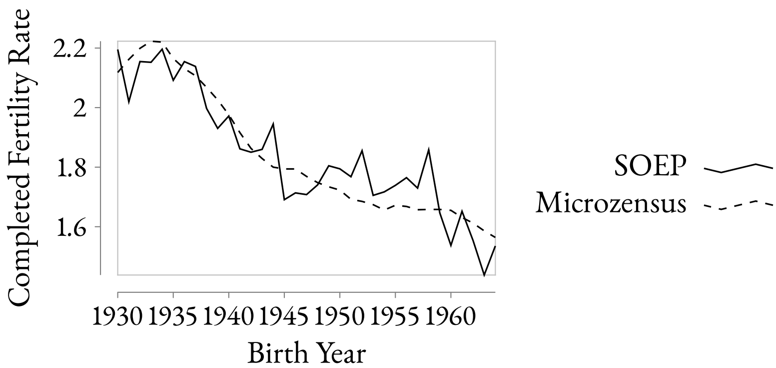
During the same time, there was a large expansion in education opportunities for women, leading to a strong increase in education levels. The education of a woman is generally negatively related to the number of children (see Baudin et al., 2015; Rindfuss et al., 1980). The educational attainment of women from the cohorts for which I have completed fertility rates are reported in Figure 3.2. The share of women with more than lower secondary education increased from 20% for the birth cohorts of 1930s to around 80% for the ones of the 1960s. If I decompose the completed fertility rate by education, as displayed in Figure 3.2, one can see that women with higher education have on average lower fertility rates, but the different education groups follow a common negative trend in the completed fertility rate. Keeping the education composition constant from the 1930, one can still see a 20% decline in the fertility rate (from 2.13 to 1.54). Thus the educational expansion alone explains only about 25% of the decline in fertility.

Furthermore, I can decompose the completed fertility rate into the extensive and the intensive fertility rate. The extensive fertility rate gives us the share of women that have at least one child. I report its inverse, the share of childless women, in Figure 3.3. The different education levels follow roughly the same trend. The share of childless women is roughly flat and around 15% until the birth cohorts of 1950 and strongly increasing afterwords to around 30% for the last observed cohort. The intensive fertility rate reports the average number of children of a

[6]Taking the average CFR in 1930 to 1934 and the average CFR from 1960 to 1964 from the SOEP data.

[7]Mikrozensus of 2016 from Statistisches Bundesamt (Destatis), retrieved in November 2017.

Figure 3.1.: *Completed Fertility*



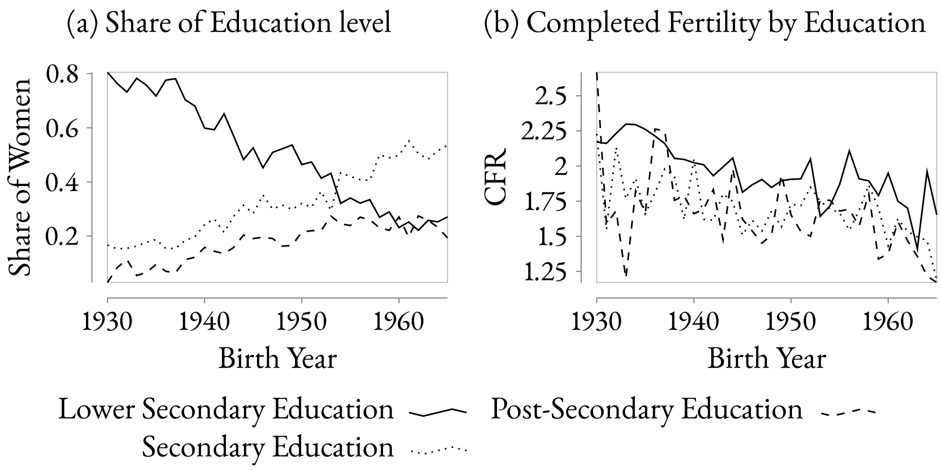
NOTE: Completed fertility rate: the average number of children of a women of a particular birth cohort. Data source: Microzensus, SOEP.

woman that has a at least one child and is displayed in Figure 3.3. The intensive fertility rate for all education groups decreased strongly for the 1930 to 1950 birth cohorts, on average from 2.6 to around 2, and stayed constant after that.

I estimate the contribution of intensive and extensive fertility to the decline in the completed fertility rate by calculating counterfactuals where I keep the extensive fertility constant and only allow the intensive fertility rate to change and vice versa. If the intensive fertility rate would not have changed, the completed fertility rate would still have decreased by 20%. Thus the decline in the extensive fertility rate explains about 70% of the total decline in fertility and the decline in intensive fertility 30%. One can clearly see the two episodes of the decline in fertility as explained above: keeping the extensive fertility rate constant would actually increase the decline in the completed fertility rate between cohorts of the early 1930s and late 1940s by 15%, as extensive fertility has actually increased during this time and mitigated the effect of the decrease in intensive fertility. The opposite is true between the cohorts born in the late 1940s and the cohorts born in the earls 1960s. Keeping the extensive fertility constant during this episode would have lead to an increase in completed fertility rate. Thus one can say that the decline in the completed fertility rate between the early 1930s and the late 1940s birth cohorts was entirely due to the decrease in intensive fertility, whereas the decline in fertility since the late 1940 birth cohorts was entirely due to the decrease in extensive fertility.

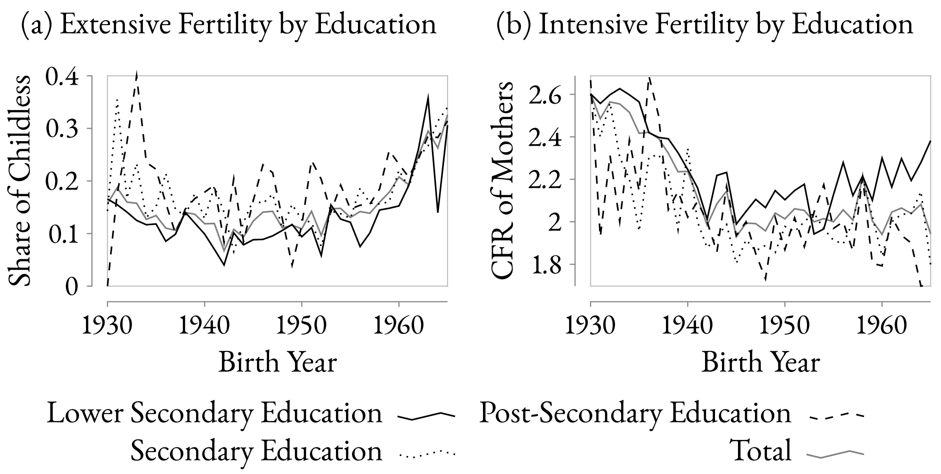
When looking at the latest cohorts for which there is data on the completed fertility displayed in Table 3.1, one can see that there are no difference between education groups in terms of extensive fertility. The share of women having at least on child is about three quarters for all education groups. The differences between

Figure 3.2.: *Education Decomposition of Fertility Rate*



Note: Share of women of a particular cohort with lower secondary, secondary, and post-secondary education. The completed fertility rate is the average number of children of a women of a birth cohort and education groups. Data source: SOEP.

Figure 3.3.: *Extensive and Intensive Fertility*



Note: Extensive and intensive fertility rate by birth cohort and education groups. Extensive fertility rate: share of childless women of a cohort. Intensive fertility rate: average number of children of a women of a cohort with at least one child. Data source: SOEP.

Table 3.1.: *Fertility Of Last Completed Cohort by Education*

	Observations	CFR	Extensive	Intensive	Age at First Birth
Lower Secondary	273.00	1.69	0.75	2.25	23.39
Secondary	567.00	1.48	0.74	2.01	25.11
Tertiary	252.00	1.38	0.74	1.88	27.12
Total	1092.00	1.51	0.74	2.04	25.14

NOTE: Completed fertility rate (CFR), extensive fertility, intensive fertility, and age at first birth by education group for cohort of women born between 1960 and 1965. Data source: SOEP.

education groups are entirely due to intensive fertility. Whereas women with at most lower secondary education and at least on child have on average 2.25 children, women with tertiary education and at least one child have on average only 1.88 children. This translates into a completed fertility rate of 1.69 for women with lower secondary education and of 1.38 for women with tertiary education. This is in contrast to the findings of Baudin et al. (2015) for the U. S., where extensive fertility follows a U-shape with regard to the education of women.

Next I investigate the opportunity cost of childbirth for women in terms of work, in particular I estimate the child penalty (the decrease in labour income after childbirth).

3.2.2. The Child Penalty

In order to understand the decision of women to have children, I consider the child penalty, the decrease in labour income after childbirth. I consider this as a measure of the opportunity cost of having a child in terms of work. In order to estimate the child penalty, I follow Kleven et al. (2018) and use an event study approach. I estimate the effect of an event (childbirth) on the income of a women and men. I estimate the effect of childbirth for 5 years prior to birth of the first child and 10 years after the birth of the first child separately for men and women. For that I restrict the panel to men and women for which I have full observations during the 5 years prior and 10 years after giving birth to their first child. I also limit the data set to these 16 years for each parent. By doing so, I get 495 women and 491men for which I observe the entire period before and after the birth of their first child. Following Kleven et al. (2018), I include fixed effects for the year of observation and the age of the parent, and I include fixed effects for the years 5 to 2 prior childbirth, the year of childbirth, and the 10 subsequent years. Thus, the terms of the leads and lags capture the variable of interest relative to the year prior childbirth. I denote t as the event time relative to childbirth, $t = 0$, a the age, and y the year. I regress the income Y_{ayt} of an individual with age a in year y and in period t relative to childbirth on the following equation:

$$Y_{ayt} = \sum_{j \neq -1} \alpha_j \mathrm{I}_{j=t} + \sum_{k} \beta_k \mathrm{I}_{k=a} + \sum_{l} \gamma_l \mathrm{I}_{l=y} + \varepsilon.$$

The first term on the left hand side is the event time fixed effects, the second term the age fixed effects, and the third term the year fixed effects. I perform this regression on net labour income, labour market participation, hours worked (conditional on labour market participation), and hourly wage (also conditional on labour market participation).

In order to illustrate the results, I scale the α_t vectors in the following way: $\tilde{\alpha}_t = \hat{\alpha}_t / E\left[\tilde{Y}_{ayt} | t\right]$ where

$$\tilde{Y}_{ayt} = \sum_{k} \hat{\beta}_k \mathrm{I}_{k=a} + \sum_{l} \hat{\gamma}_l \mathrm{I}_{l=y}$$

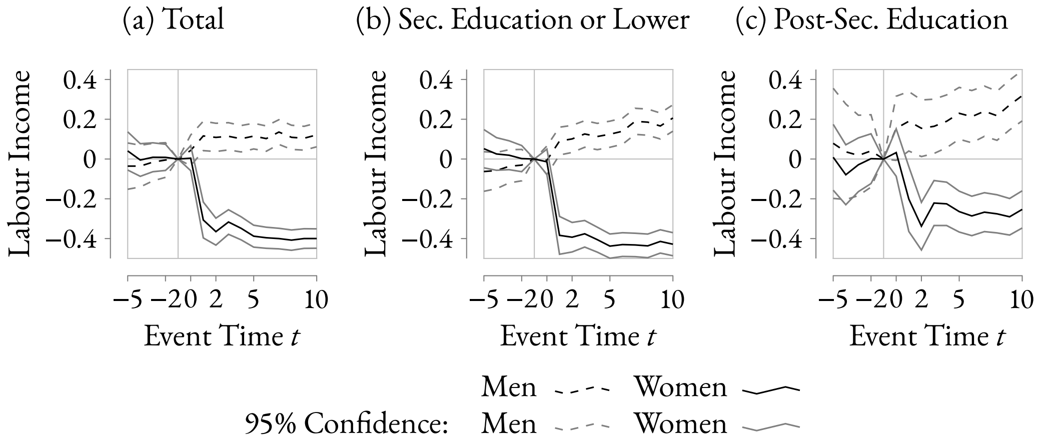
is the predicted income of an individual in period t. This implies that $\tilde{\alpha}_t$ is the impact of childbirth on income in event time t relative to the income in event time $t = -1$.

The impact of childbirth on women's and men's net labour income is shown in Figure 3.4. One can see that there is a strong drop in women's labour income after childbirth, a drop that is not present for men. This is consistent with the findings of Kleven et al. (2018). The drop is very similar for women with only secondary education or lower and women with post-secondary education. For women with secondary education or lower, the net labour income decreases by around 40%, and for women with post-secondary education by around 30%. The average decrease in net labour income is around 40%. For all education groups, this drop in labour market income is persistent over the entire observed 10 years. As can be seen in Figure 3.5 and Figure 3.6, this is almost entirely due to a drop in labour market participation, whereas the hourly wage stays completely flat (Figure 3.7).[8] For men, there is no negative change in any of the measured indicators.

The child penalty as estimated above does only compare the income and labour

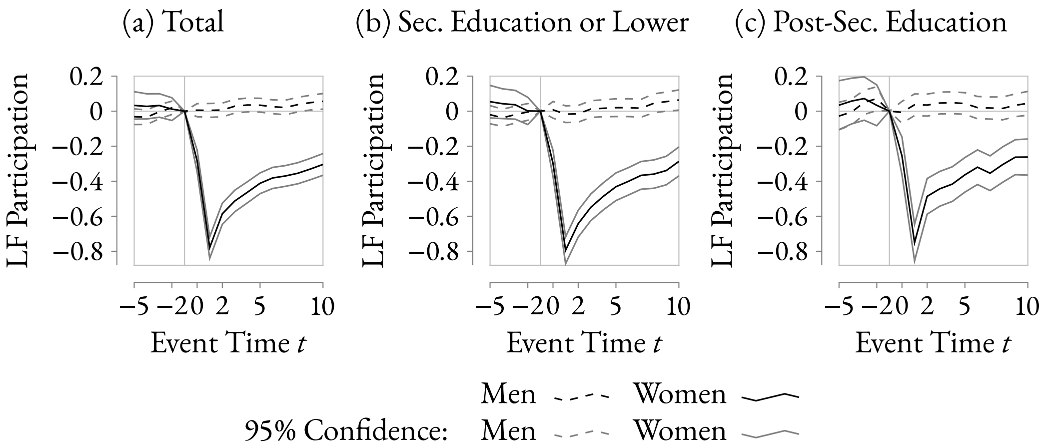
[8] The hourly wage has a high variance, as after childbirth the participation is very low and thus the number of observations to estimate the effects on is very small.

Figure 3.4.: *Child Penalty Net Labour Income*



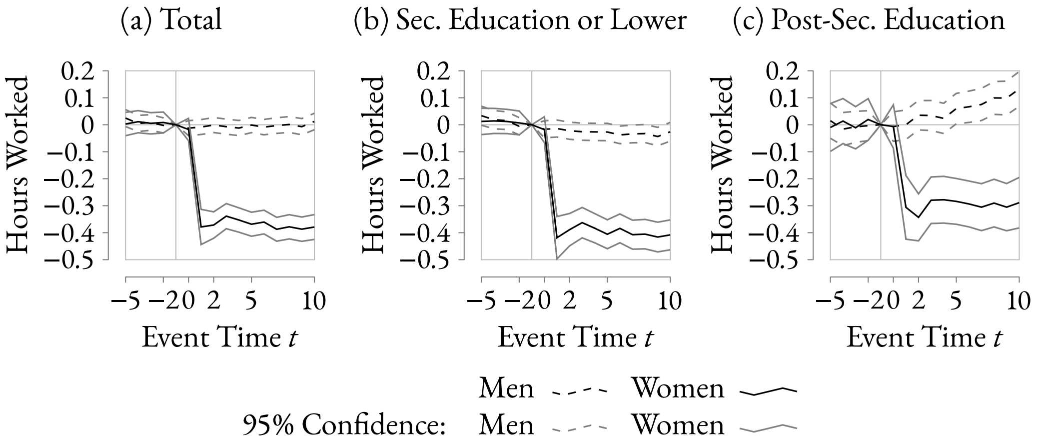
Note: Net labour income relative to $t-1$, the year before the birth of the first child. Data source: SOEP.

Figure 3.5.: *Child Penalty Participation*



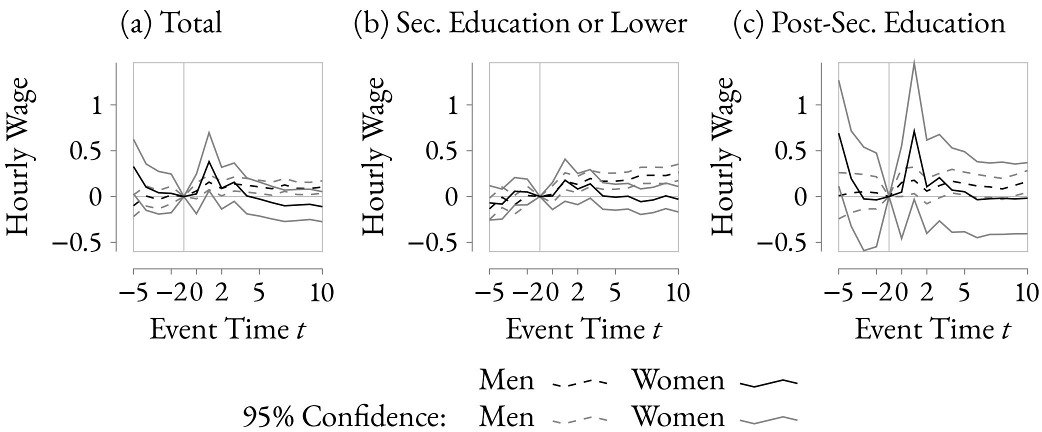
Note: Labour force participation relative to $t-1$, the year before the birth of the first child. Data source: SOEP.

Figure 3.6.: *Child Penalty Hours Worked*



NOTE: Hours worked conditional on labour force participation relative to $t-1$, the year before the birth of the first child. Data source: SOEP.

Figure 3.7.: *Child Penalty Net Hourly Wage*



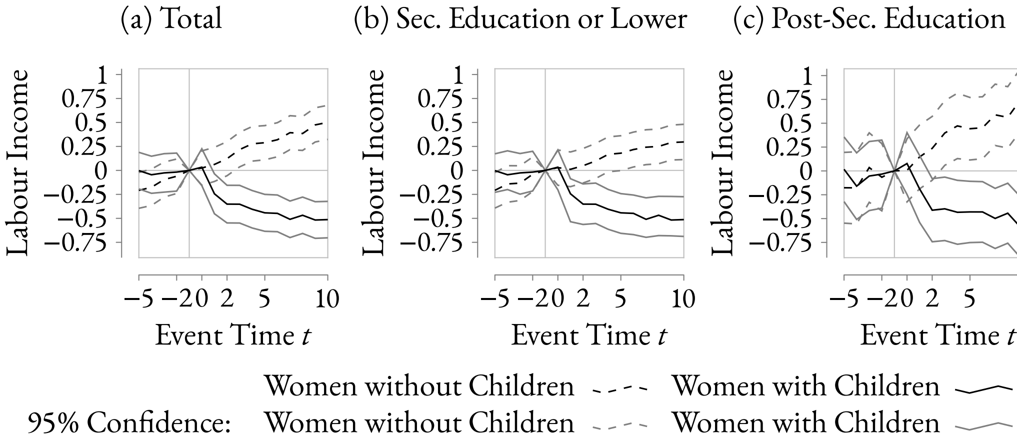
NOTE: Hourly wage conditional on labour force participation relative to $t-1$, the year before the birth of the first child. Data source: SOEP.

indicators of women that give birth before and after their first child. Hence, this measure gives us not a comparison of what the income would have been if the woman would not have had a child. This means that it potentially underestimates the child penalty because of sorting of women with strong preferences for children into occupations with lower child penalties (Adda et al., 2017). In order to check for this, I follow Kleven et al. (2018) and also perform a difference in difference estimation of women that have a child and women that do not have children. For doing so, I only consider women where the last observation is at an age above 40, as this is considered to normally be the oldest age for a first child. I estimate the distribution of childbirth over women's age, and draw placebo births for the childless women from this estimation. Then I perform a difference in difference estimation along the lines of the event study approach comparing the two groups. A detailed description of the approach as well as the results for all variables can be found in Appendix B.1.

Figure 3.8 shows the result on the difference-in-difference estimation on net labour income. One can see that for both women with at most secondary education as well as for women with post-secondary education, there is a child penalty of around 50% compared to their income in $t-1$. The comparison groups of women with no children shows an increase in income after childbirth though, leading to a higher overall child penalty than the one estimated by the event study including only women that evetually become mothers. The results also point towards a higher child penalty for women with higher education, which is in line with Adda et al. (2017). Unfortunately the sample for the difference-in-difference approach is not big enough to find any significant effects for the other variables.

This analysis shows that there is a substantial opportunity cost in terms of women's work time, leading to a child penalty in women's labour income. I think that due to the increase in women's labour market opportunities, this opportunity costs was becoming increasingly important, which could explain the decline in completed and intensive fertility, and the increase in childlessness. To examine this hypothesis, I develop a model featuring a per child time costs for women in the next section, and show in a calibration exercise in Section 3.4 that this model can generate the fertility patterns shown above.

Figure 3.8.: *Child Penalty Women Net Labour Income – Diff-In-Diff*



NOTE: Net labour income of women relative to $t-1$, the year before the birth of the first child. For childless women, placebo births are assigned according to distribution of age at first birth of women with children. For a description of the method see Appendix B.1. Data source: SOEP.

3.3. The Model

I develop an OLG model with endogenous fertility of an economy where households are composed of one woman and one man. The agents live for three periods. In the first period, agents are children and receive education, in the second period they are adults, form a couple and make decision about how much to work, how many children to have, and how to educate them. In the third period they retire and consume their savings. The decisions are made jointly at the household level and the households are heterogeneous with regard to their preference for children. They inhabit a small open economy with two types of labour: physical and mental. Women can only supply up to one unit of mental labour, whereas men supply both one unit of mental and one unit of physical labour.

3.3.1. The Economy

The economy is a small open economy that evolves according to a Galor and Weil (1996) type production function, where man can provide both physical and mental labour and women only mental labour.[9] The production function of this econ-

[9]Differences in wages come from a wide range of reasons, for a discussion of the gender wage gap see Kleven et al. (2018). For simplicity I assume here that they only stem from differences in types of supplied labour, but the main results of this paper hold up for different causes for the gender wage gap.

omy is

$$Y_t = a\left[\alpha K_t^{\varrho} + (1-\alpha)M_t^{\varrho}\right]^{\frac{1}{\varrho}} + bP_t,$$

where K_t is the capital, M_t is the mental labour, and P_t the physical labour in this economy in period t. I assume that $a, b > 0$, $\alpha \in (0,1)$, and $\varrho \in (-\infty, 1)$. Following Galor and Weil (1996), I write the production function in per household terms:

$$y_t = a\left[\alpha k_t^{\varrho} + (1-\alpha)m_t^{\varrho}\right]^{\frac{1}{\varrho}} + bp_t,$$

where y_t is the per household output, k_t is the per household level of capital, m_t is the per household supply of mental labour, and p_t is the per household supply of physical labour. Men can supply both one unit of mental and one unit of physical labour, and women only one unit of mental labour. Maximising the production function with respect to labour, I get the following relationships between wages on physical and mental labour $w_{p,t}$ and $w_{m,t}$, respectively, and the per household levels of mental labour supply and capital:

$$w_{p,t} = b,$$

and

$$w_{m,t} = a(1-\alpha)m_t^{\varrho-1} \cdot \left[\alpha k_t^{\varrho} + (1-\alpha)m_t^{\varrho}\right]^{\frac{1-\varrho}{\varrho}}. \tag{3.1}$$

I assume that the economy is a small open economy with the interest rate r given exogenously. By optimising the production function with respect to capital, I get the optimal amount of capital given the labour supply m_t:

$$k_t^* = \frac{(a\alpha)^{\frac{\varrho}{1-\varrho}} m_t (1-\alpha)^{\frac{1}{\varrho}}}{\left[r^{\frac{\varrho}{1-\varrho}} - \alpha(a\alpha)^{\frac{\varrho}{1-\varrho}}\right]^{\frac{1}{\varrho}}}.$$

One can insert this into (3.1) to get the equilibrium wage on mental labour

$$w_{m,t} = a(1-\alpha)^{\frac{1}{\varrho}} r \left[r^{\frac{\varrho}{1-\varrho}} - \alpha(a\alpha)^{\frac{\varrho}{1-\varrho}}\right]^{\frac{\varrho-1}{\varrho}}.$$

This wage only depends on parameters representing the productivity of the economy and is independent of the supply of mental labour m_t. The supply of mental labour is $m_t \in [0, 2]$ and depends on the wages in this economy. As men supply both physical and mental labour, the wage for men is $w_{m,t} + w_{p,t}$ and the wage for women (who only supply mental labour) is equal to $w_{m,t}$. Thus in this setting, women always earn a lower wage than men. I will show below that the labour supply of men is inelastic, and as each household comprises of one man an one women, the per household supply of physical labour is $p_t = 1$ and the supply of mental labour is $m_t \in [1, 2]$. Next we derive the decision on consumption, education, and fertility based on these wages.

3.3.2. The Private Decision

In period t the households i decide how much both members work and how many children to have, as well as how much to invest into their children's education. This decisions are based on the following utility function over the number of children $n_{i,t}$, the children's education $h_{i,t}$, and consumption $\{c_{i,t}, c_{i,t+1}\}$:

$$U_i(n_{i,t}, c_{i,t}, c_{i,t+1}, h_{i,t}) = \log(c_{i,t}) + \beta \log(c_{i,t+1}) + \gamma \log(n_{i,t}h_{i,t} + v_i), \tag{3.2}$$

where β is the discount factor on future consumption, γ captures the altruism of the adults towards their children, and v_i is a preference parameter that allows for zero fertility as in Baudin et al. (2015). The households are heterogeneous with regard to this v_i. Their decisions are subject to the following budget constraint

$$c_{i,t} + n_{i,t}x_{i,t} + s_{i,t} = w_m l_{w,i,t} + (w_{m,t} + w_{p,t})l_{m,i,t}, \tag{3.3}$$

$$c_{t+1} = (1 + r)s_t, \tag{3.4}$$

where $l_{w,i,t}$ and $l_{m,i,t}$ are the labour supply of woman and men respectively. $x_{i,t}$ is the education spending per child by the parents, and $s_{i,t}$ are the savings for second period consumption.

Additional to the budget constraint, household face a time constraint. The household has to spend time on home-production. Childrearing has a per-child cost of $\vartheta \in (0, 1)$. Additionally each household has to perform a fixed amount of home-production $\Psi \in [0, 1)$. The home-production performed by the women is $e_{w,i,t}$ and the one performed by the man is $e_{m,i,t}$. I have thus the following condition for time spend on home-production:

$$\Psi + n_{i,t}\vartheta = e_{w,i,t} + e_{m,i,t}.$$

Each spouse has a total time normalised to 1, which they can spend on working, childrearing, and home production. The total time constraints for each spouse is

$$1 \geq l_{w,i,t} + e_{w,i,t},$$

and

$$1 \geq l_{m,i,t} + e_{m,i,t}.$$

As women earn a smaller wage in this economy than men, and income is linear in hours worked, the home production and childrearing is done entirely by the women.[10] This leads to

$$l_{m,t} = 1, \tag{3.5}$$

$$l_{w,t} = 1 - e_{w,i,t} = 1 - \Psi - n_{i,t}\vartheta. \tag{3.6}$$

The human capital of the children $h_{i,t}$ depends on education spending $x_{i,t}$ with which the parents pay teachers that receive a wage on mental labour $w_{m,t}$. Human capital is produced according

$$h(x_{i,t}) = \left(\frac{x_{i,t}}{w_{m,t}}\right)^{\eta}, \tag{3.7}$$

where $\eta \in (0, 1)$ is a parameter signifying the efficiency of human capital formation.

One can now find the optimal number of children and private education spending conditional on $n_{i,t} > 0$ by maximising (3.2) subject to (3.3), (3.4), (3.7), (3.5), and (3.6). The optimal values of $\{c_{i,t}, c_{i,t+1}, n_{i,t}, x_{i,t}\}$ are equal to

[10] I abstract from the case where the household choose an $n_{i,t} > (1-\Psi)/\vartheta$, i. e. where men would also have to spend time on childrearing. In this specification, in the case of $n_{i,t} > (1 - \Psi)/\vartheta$, the households hire a household help earning a wage on mental labour $w_{m,t}$. In Section 3.4, I estimate ϑ to be equal to 0.11 and Ψ to be equal to 0.38, meaning that this case would only apply for women with more than 6 children, which is a small minority.

$$
\begin{aligned}
c_{i,t}^{*} &= \frac{1}{1+\beta+\gamma}\left[(w_{m,t}+w_{p,t})+w_{m,t}(1-\Psi)+\frac{w_{m,t}v_i}{\eta^{\eta}(1-\eta)^{1-\eta}}\vartheta^{1-\eta}\right],\\
c_{i,t+1}^{*} &= \frac{\beta(1+r)}{1+\beta+\gamma}\left[(w_{m,t}+w_{p,t})+w_{m,t}(1-\Psi)+\frac{w_{m,t}v_i}{\eta^{\eta}(1-\eta)^{1-\eta}}\vartheta^{1-\eta}\right],\\
n_{i,t}^{*} &= \frac{\gamma(1-\eta)}{1+\beta+\gamma}\frac{\frac{w_{m,t}+w_{p,t}}{w_{m,t}}+(1-\Psi)}{\vartheta}-\frac{1+\beta}{1+\beta+\gamma}\frac{(1-\eta)^{\eta}}{\eta^{\eta}}\frac{v_i}{\vartheta^{\eta}},
\end{aligned}
\tag{3.8}
$$

and

$$
x_{i,t}^{*} = \frac{\eta\vartheta}{1-\eta}w_{m,t}.
$$

The households cannot have a negative number of children, therefor there is a zero lower bound for $n_{i,t}$ with $n_{i,t} \geq 0$. Parents with an $n_{i,t}^{*} \leq 0$ prefer to be childless than to have a child. I derive the value $\hat{v}_t$ of v_i for which a household is indifferent between having a child or not by solving (3.8) for the v_i for which $n_{i,t}^{*} = 0$:

$$
\hat{v}_t = \frac{\gamma}{1+\beta}\frac{\eta^{\eta}(1-\eta)^{1-\eta}}{\vartheta^{1-\eta}}\left[\frac{w_{m,t}+w_{p,t}}{w_{m,t}}+(1-\Psi)\right].
$$

If the household has a $v_i < \hat{v}_t$, it stays childless and optimal consumption becomes

$$
\bar{c}_{i,t}^{*} = \frac{1}{1+\beta}\left[(w_{m,t}+w_{p,t})+w_{m,t}(1-\Psi)\right]
$$

and

$$
\bar{c}_{i,t+1}^{*} = \frac{\beta(1+r)}{1+\beta}\left[(w_{m,t}+w_{p,t})+w_{m,t}(1-\Psi)\right]
$$

with $n_{i,t} = 0$ and $x_{i,t} = 0$.

Next I am going to look at the effect of changes in wages on the equilibrium level of fertility at the household level, i. e. how $n_{i,t}^{*}$ and $\hat{v}_t$ are affected by changes in $w_{m,t}$ and $w_{p,t}$.

3.3.3. The Effect of Wage Changes on Fertility

First I investigate the effect of a change in the wage on physical labour $w_{p,t}$ (a change in b). An increase in the wage on physical labour $w_{p,t}$ has no effect on the education spending, but increases the number of children per household:

Proposition 7. *An increase in the wage on physical labour $w_{p,t}$ has no effect on $x^*_{i,t}$ but increases the number of children $n^*_{i,t}$ as well as $\hat{v}_t$*

Proof. The first derivative of $x^*_{i,t}$ with respect to $w_{p,t}$ is 0, and the first derivative of $n^*_{t,i}$ with respect to $w_{p,t}$ is:

$$\frac{\partial n^*_{i,t}}{\partial w_{p,t}} = \frac{\gamma(1-\eta)}{1+\beta+\gamma}\frac{1}{\vartheta}\frac{1}{w_{m,t}}, \tag{3.9}$$

which is positive. The first derivative of $\hat{v}_t$ with respect to $w_{p,t}$ is equal to

$$\frac{\partial \hat{v}_t}{\partial w_{p,t}} = \frac{\gamma}{(1+\beta)}\frac{\eta^\eta(1-\eta)^{1-\eta}}{\vartheta^{1-\eta}}\frac{1}{w_{m,t}}, \tag{3.10}$$

which is also positive. □

Thus a change in the wage on physical labour does not trigger a quantity/quality trade off in this model (i. e. it does not lead to an increase in the investment into human capital and a decrease in the number of children), but increases the number of children in a particular household. It also increases the value of $\hat{v}_t$ below which households stay childless, i. e. leads to less households being childless. This is because in this model, an increase in the wage on physical labour makes investment into education relatively cheaper, as the wages of teachers and the opportunity cost of children (the women's wage) do not change, but the total income of a household increases. Thus there is an income effect leading to higher fertility, but no substitution effect leading to lower fertility and higher education investment.

When looking at the effect of a change in the wage on mental labour (and thus in the gender wage gap), one can see that an increase in the wage on mental labour (i. e. due to an increase in a) increases per child spending on private education $x^*_{i,t}$, and decreases the number of children $n^*_{i,t}$ as well as the marginal value of $\hat{v}_t$ for which a household is indifferent between having children or not.

Proposition 8. *An increase in the wage on mental labour $w_{m,t}$ increases the per child spending in private education $x^*_{i,t}$, and decreases the number of children $n^*_{i,t}$ as well as $\hat{v}_t$.*

Proof. The first derivative of $x^*_{i,t}$ with repect to $w_{m,t}$ is

$$\frac{\partial x^*_{i,t}}{\partial w_{m,t}} = \frac{\eta\vartheta}{1-\eta} w_{m,t}$$

which is positive. The first derivative of $n^*_{i,t}$ with repect to $w_{i,m}$ is

$$\frac{\partial n^*_{i,t}}{\partial w_{m,t}} = -\frac{\gamma(1-\eta)}{(1+\beta+\gamma)} \frac{1}{\vartheta} \frac{w_{p,t}}{w^2_{m,t}}$$

which is negative. The first derivative of $\hat{v}_t$ with repect to $w_{m,t}$ is

$$\frac{\partial \hat{v}_t}{\partial w_{m,t}} = -\frac{\gamma}{(1+\beta)} \frac{\eta^\eta(1-\eta)^{1-\eta}}{\vartheta^{1-\eta}} \frac{w_{p,t}}{w^2_{m,t}}$$

which is also negative. □

The substitution effect outweighs here the income effect leading to higher spending on education and less fertility. Additionally, an increase in the wage on mental labour decreases the threshold level of $\hat{v}_t$ above which a household chooses not to have children. Thus there is on the household level a quantity/quality trade off with respect to the wage on mental labour $w_{m,t}$, but not with respect to the wages on physical labour.

These effects are considered at the household level thus far, next I look at the effect of changes in wages on the aggregate level of labour supply, fertility, and childlessness.

3.3.4. Aggregate Fertility

I assume that households are homogeneous except to v_i and that v_i is distributed according to a log normal distribution function, i. e. that $v \sim LN(\mu, \sigma)$, where μ and σ are the mean and the standard deviation of the underlying normal distribution.

As the individual labour supply of men is inelastic with respect to the wages, the aggregate labour supply of men $L_{m,t}$ is always equal to 1. By aggregating the individual labour supply of women $l_{w,i,t}$ over the distribution of v I get the aggregate labour supply of women $L_{w,t}$ of the economy corresponding to the wages $w_{m,t}$ and $w_{p,t}$:

$$L_{w,t} = \int_0^{\hat{v}_t} (1 - \Psi - n_t^*(v_i)\vartheta)g(v_i)dv_i + \int_{\hat{v}_t}^{\infty} (1 - \Psi)g(v_i)dv_i,$$

where $g(v_i)$ is the probability density function of v_i. Using (3.8), one can rewrite this as

$$\begin{aligned} L_{w,t} = (1-\Psi) &- \frac{\gamma(1-\eta)}{1+\beta+\gamma}\left[\frac{w_{m,t}+w_{p,t}}{w_{m,t}} + (1-\Psi)\right] G(\hat{v}_t) \\ &+ \left[\frac{1+\beta}{1+\beta+\gamma}\frac{(1-\eta)^{\eta}}{\eta^{\eta}}\frac{1}{\vartheta^{\eta-1}}\right]\int_0^{\hat{v}_t} v_i g(v_i)dv_i, \end{aligned} \tag{3.11}$$

where $G(v_i)$ is the cumulative density function of v_i. I insert for $G(\hat{v}_t)$ and $\int_0^{\hat{v}_t} v_i g(v_i)dv_i$ the cumulative density function and the partial expectation of the lognormal distribution to get

$$\begin{aligned} L_{w,t} = (1-\Psi) &- \frac{\gamma(1-\eta)}{1+\beta+\gamma}\left[\frac{w_{m,t}+w_{p,t}}{w_{m,t}} + (1-\Psi)\right] \Phi\left(\frac{log(\hat{v}_t)-\mu}{\sigma}\right) \\ &+ \left[\frac{1+\beta}{1+\beta+\gamma}\frac{(1-\eta)^{\eta}}{\eta^{\eta}}\frac{1}{\vartheta^{\eta-1}}\right] e^{\mu+\frac{\sigma^2}{2}}\Phi\left(\frac{log(\hat{v}_t)-\mu-\sigma^2}{\sigma}\right), \end{aligned}$$

where Φ is the cumulative density function of the $N(0,1)$ distribution.

In the same way I can derive the equilibrium completed fertility rate of the economy, which is equal to

$$N_t = \int_0^{\hat{v}_t} n_t^*(v_i)g(v_i)dv_i,$$

which one can rewrite as

$$N_t = \frac{\gamma(1-\eta)}{1+\beta+\gamma}\frac{\frac{w_{m,t}+w_{p,t}}{w_{m,t}} + (1-\Psi)}{\vartheta}G(\hat{v}_t) - \left[\frac{1+\beta}{1+\beta+\gamma}\frac{(1-\eta)^{\eta}}{\eta^{\eta}}\frac{1}{\vartheta^{\eta}}\right]\int_0^{\hat{v}_t} v_i g(v_i)dv_i.$$

By again inserting the cumulative density function and the partial expectation of the lognormal distribution, I get

$$N_t = \frac{\gamma(1-\eta)}{1+\beta+\gamma}\frac{\frac{w_{m,t}+w_{p,t}}{w_{m,t}}+(1-\Psi)}{\vartheta}\Phi\left(\frac{log(\hat{v}_t)-\mu}{\sigma}\right)$$
$$-\left[\frac{1+\beta}{1+\beta+\gamma}\frac{(1-\eta)^\eta}{\eta^\eta}\frac{1}{\vartheta^\eta}\right]e^{\mu+\frac{\sigma^2}{2}}\Phi\left(\frac{log(\hat{v}_t)-\mu-\sigma^2}{\sigma}\right)$$

The equilibrium extensive fertility rate of this economy is equal to

$$E_t = G(\hat{v}_t) = \Phi\left(\frac{log(\hat{v}_t)-\mu}{\sigma}\right)$$

and the equilibrium intensive fertility rate $I_t = N_t/E_t$ is equal to

$$I_t = \frac{\gamma(1-\eta)}{1+\beta+\gamma}\frac{\frac{w_{m,t}+w_{p,t}}{w_{m,t}}+(1-\Psi)}{\vartheta}$$
$$-\left[\frac{1+\beta}{1+\beta+\gamma}\frac{(1-\eta)^\eta}{\eta^\eta}\frac{1}{\vartheta^\eta}\right]e^{\mu+\frac{\sigma^2}{2}}\frac{\Phi\left(\frac{log(\hat{v}_t)-\mu-\sigma^2}{\sigma}\right)}{\Phi\left(\frac{log(\hat{v}_t)-\mu}{\sigma}\right)}.$$

Using this expressions for labour supply, fertility, and childlessness, I investigate how changes in the wage on mental labour affect the economy's fertility rates.

Proposition 9. *An increase in $w_{m,t}$ decreases the total fertility rate N_t and the extensive fertility rate E_t, and increases female labour force participation $L_{w,t}$.*

Proof. When considering the effect of $w_{m,t}$ on N_t, I find that an increase in $w_{m,t}$ decreases $\hat{v}_t$ and decreases $n_t^*(v_i)$ for any given v_i (Proposition 8), thus an increase in $w_{m,t}$ decreases N_t. It follows directly from the fact that N_t is decreasing that L_t is increasing. The first derivate of E_t with respect to $w_{m,t}$ is

$$\frac{\partial E_t}{\partial w_{m,t}} = -\phi\left(\frac{\log(\hat{v}_t)-\mu}{\sigma}\right)\frac{1}{\sigma}\frac{w_{p,t}}{w_{m,t}(w_{m,t}+w_{p,t})+w_{m,t}^2(1-\Psi)},$$

where ϕ is the probability density function of the $N(0,1)$ distribution. This is always negative. □

Thus an increase in the wage on mental labour decreases the total fertility rate by increasing the opportunity cost of children in terms of women's wages. Less women decide to have children, and those who do have less children. The effect on the intensive fertility rate is not obvious though, as $I_t = N_t/E_t$, and both N_t and E_t are decreasing in $w_{m,t}$. If intensive fertility is decreasing or not depends on whether N_t or E_t decreases faster.

This comparative static analysis shows that an increase in the wage on mental labour (a decrease in the gender wage gap) leads to a decrease in fertility, and an increase in childlessness. In the next section I will calibrate the model and simulate values of $L_{w,t}$, N_t, E_t, and I_t to see if the model can generate the fertility patterns observed in the data.

3.4. Calibration

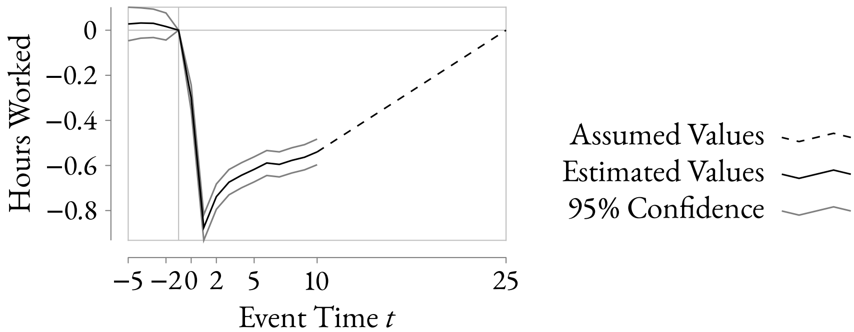
To assess the validity of the model, I calibrate the parameters of the model to show that it can replicate the patterns in completed fertility observed in Section 3.2. I choose the parameters to match the data of the last observed cohort (early 1960s), and then show that with a change in the relative wage of women (the relative wage on mental labour), I can generate the historical patterns in fertility.

I do not have a historical series on the gender wage gap by which this model is driven, but I have data on the labour force participation rate of women, which I will use as an alternative target. The model is set in terms of hours worked, for which I have no long run series, therefore I use the female labour force participation rate as a proxy. First, I choose parameter values that fit the data of the last completed cohort (early 1960s), and then change the value of the wage premium of men (the inverse of the gender wage gap) in order to match the female labour force participation rate observed in the data. For the labour force participation in Germany I use data from the U. S. Bureau of Labor Statistics (BLS) measured 40 years after birth of the cohort (as there is no additional motherhood after an age of 40).[11]

The parameter values used in the calibration are displayed in Table 3.2. I choose standard values for the discount factor β and the interest rate r of 0.95 and 0.05, respectively. For Ψ I estimate the average time that women from the age of 20 to 65 spend on household work as share of non-leisure time from NTTA data, which is equal to 0.38 (they spend on average 4 hours on household work, which is 38%

[11]U.S. Bureau of Labor Statistics, Labor Force Participation Rate for Women in Germany (discontinued) [DEULFPWNA], retrieved from FRED, Federal Reserve Bank of St. Louis; (https://fred.stlouisfed.org/series/DEULFPWNA), March 25, 2019.

Figure 3.9.: *Child Penalty in Total Hours Worked*



NOTE: Hours worked unadjusted for labour force participation relative to $t-1$, the year before the birth of the first child. Estimated total child penalty 10.92, per year during this 25 years 0.42. This amounts to a child penalty in work time of 0.23 as share of a 45 year work life. This amounts to a per child value of 0.11 of total working time assuming an intensive fertility rate of 2.04 as in the last completed cohort (early 1960s). Data source: SOEP.

of 10.6 hour of non-leisure work time per day – 24h day discounting 8 hours of sleep and the average time devoted to leisure activity of 5.4h).[12] For an initial value of $w_m/(w_m + w_p)$, I use the gender wage gap of 2006 from Eurostat of 0.77, which results in a value for the male wage premium of 1.3.[13]

I estimate the per child penalty for women in terms of working hours ϑ by using the event study approach from Section 3.2.2. For this, I estimate the child penalty in total hours worked (taking into account labour force participation). I do only observe the child penalty for the first 10 years after childbirth, so I assume that the child penalty will reverse to 0 after 15 years. This is displayed in Figure 3.9. Giving the observed conversion back to 0 in the first 10 years, I think that this assumption is fairly conservative and likely underestimates the total child penalty.

I find a total child penalty of 10.92, which is equivalent to an average yearly decline in total hours worked of 0.42. Assuming a career of 45 years, this implies a child penalty in terms of life-time work of 0.23. As the women in this sample have not necessarily completed fertility, I divide this by the number of children of the last observed completed cohort (2.04) to get an estimate of the per child penalty of 0.11. The completed fertility rate is declining over the observed period and the

[12] Source Vargha, L., Šeme, A., Gál, R., I., Hammer, B., Sambt, J. (2016): European National Time Transfer Accounts. Available at: http://www.wittgensteincentre.org/ntadata.

[13] Data: Eurostat, Labour market (including Labour Force Survey): earn_grgpg gender pay gap in unadjusted form by NACE Rev. 1.1 activity - structure of earnings survey methodology, last update: 10-08-2016 (http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=earn_gr_gpg&lang=en).

Table 3.2.: *Calibration*

	Value	Target	Data	Model
β	0.95	Discount factor		
r	0.05	Interest rate.		
Ψ	0.38	Share of time devoted to household work by German women.[1]		
$\frac{w_m+w_p}{w_m}$	1.3	Gender wage gap in 2006.[2]		
ϑ	0.11	Child penalty of for ten years.[3]		
η	0.07	Female labour force participation in early 2000s.[4]	0.49	0.45
γ	0.89	Extensive fertility in early 1960s	0.74	0.74
μ	1.32	Intensive fertility in early 1960s.	2.04	2.05
σ	0.54	Standard deviation of intensive fertility in early 1960s.	1.21	1.09

Note: Parameter values chosen in calibration of the model to match cohort born in the early 1960s. Data source: SOEP.
[1] Average time that women from the age of 20 to 65 spend on household work as share of non-leisure time. Data source: NTTA.
[2] Inverse of gender pay gap. Source: Eurostat.
[3] Estimation based on event study approach, assumed reversion of child penalty after 15 years. Data source: SOEP.
[4] Source: BLS.

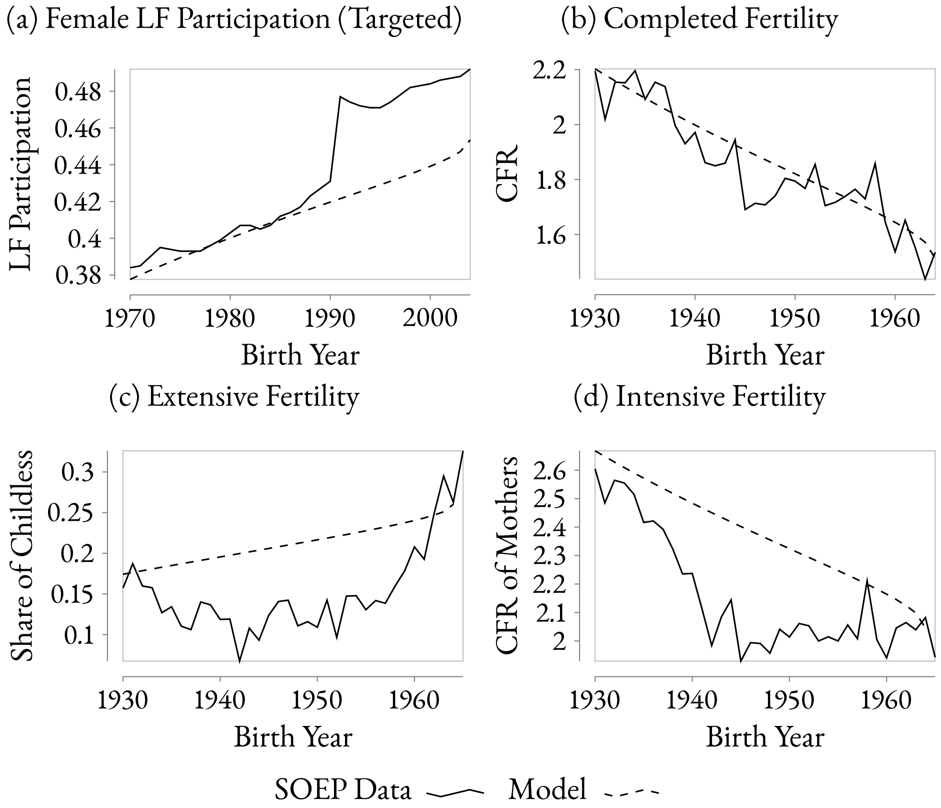
average number of children for this cohorts is probably lower, which again means that this estimated per child penalty in hours worked is likely underestimated.

The other values are chosen such that specific moments in the data about the last completed cohort (women born in 1960 to 1965) are matched. In particular, I choose γ such that extensive fertility matches the one observed in the data, μ such that intensive fertility matches the one observed in the data, and σ such that the standard deviation of intensive fertility matches the one in the data.[14] Finally I choose η such that the labour force participation matches the one observed in the early 2000s.

The model manages to match the extensive fertility rate and the intensive fertility rate almost exactly, with an extensive fertility rate of 0.74 in both the data and the model, and an intensive fertility rate of 2.04 in the data and of 2.05 in the model. The model slightly underestimates the standard deviation of intensive fertility (1.21 in the data vs. 1.09 in the model) and the model underestimates the labour force participation rate with a value in the model of 0.45 compared to a value of 0.49 in the data. I think the latter might be due to the fact that this model and the assumed values of Ψ and ϑ are based on hours worked, whereas the data are on labour force participation. As the labour force participation rate does not take into account that as an effect to childbirth women that choose to work after childbirth also reduce the hours they work, it is a lower bound to the actual role

[14] In order to estimate the standard deviation of the model, I use a Monte Carlo estimation with 10000 observations.

Figure 3.10.: *Simluation Results*



NOTE: Simulation of completed fertility rate, intensive fertility rate, and childlessness based on calibration target to to cohorts of early 1960s, and gender wage gap adjusted such that model matches female labour source participation rate. Data source: SOEP, BLS.

that childbirth plays for female labour supply.

Using this initial calibration, I now generate a series of the wage premium for men $(w_{p,t} + w_{m,t})/w_{m,t}$ that matches the pattern observed in the female labour supply.[15] The resulting series for the completed fertility rate, extensive fertility rate, and intensive fertility rate are displayed in Figure 3.10.

One can see that by choosing the relative wage of women such that the labour force participation of women matches the data, the model can generally match the pattern in observed fertility. The calibration matches very well the decline in completed fertility, and generates a decline in intensive fertility and an increase in childlessness that matches the difference between the generations born in the 1930s and the generations born in the early 1960s. However the calibration of the model

[15]In particular I keep w_p constant at 300, and change $w_{m,y}$ for the years $y \in \{1970, \dots, 2005\}$ according to $w_{m,2005-t} = 1000 - 90t^{0.5}$ with $t \in \{0, \dots, 35\}$. Note that for this calibration the actual value of $w_{m,t}$ and $w_{p,t}$ do not matter, only their relative size.

is not able to generate the increasing importance of childlessness and the decreasing role of intensive fertility for the decline in the fertility rate. Thus, this endogenous fertility model of a quantity/quality trade-off with opportunity cost of children for women in terms of time is able to generate both a decline in completed fertility and intensive fertility, as well as an increase in extensive fertility.

3.5. Conclusion

This paper investigates the role that intensive and extensive fertility play in the decline in fertility in Germany for the cohorts of women born between 1930 and 1965. I show that there were two stages: the decline in fertility for cohorts of women born before the 1950s was entirely due to a decline in intensive fertility, whereas the decline in fertility for the cohorts born after 1950 was entirely due to an increase in childlessness. I argue that the importance of childlessness is ad odds with current models of a quantity/quality trade off. Therefor I study the importance of the opportunity cost of children for women in terms of the so called child penalty using an event study approach. The decline in women's labour income after childbirth seems to be almost entirely due to a decline in hours worked.

Based on this observations, I build an overlapping generations model with a quantity/quality trade-off that allows for childlessness. I show that this model can simulate the "essential complementarity" of non-zero fertility and the quantity/quality trade-off. I show in a calibration exercise, that through a decrease in the gender wage gap the model can generate the fertility patterns observed in the data.

I think that future research should be conducted to evaluate this mechanisms for other countries as well as in earlier stages of economic development. I argue that the opportunity costs of children in terms of work time for women are essential to understand the decline in fertility and I expect family policies that are directed at these might have the best chance of affecting fertility in a positive way. Therefore I think it would be interesting to use the framework developed in this paper to analyse policies for their effect on fertility and childlessness.

4 Inequality and Education Spending in a Greying Society

With Gianko Michailidis

4.1. Introduction

Population ageing has become an issue of growing concern for OECD countries, especially as the generation of "baby boomers" reach retirement age, putting considerable pressure on pensions systems and the welfare state. Parallel to this, during the last decades there was a strong increase in income inequality. These trends have drawn attention to the public finance of education and the sustainability of public pensions as they aggravate two of the main political conflicts over the welfare state. The increase in income inequality intensifies the intragenerational conflict between rich and poor over redistribution in the form of public education. Population ageing exacerbates the intergenerational conflict over the allocation of resources between elderly and young.

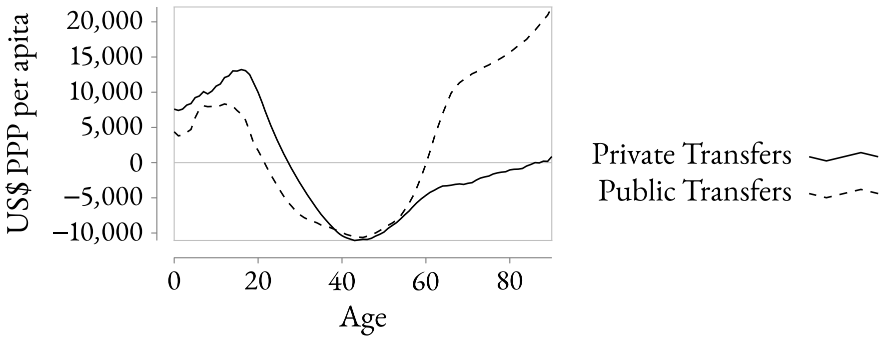
These conflicts are examined in the literature on the political economy of pensions and education. In this literature, most of the studies consider these conflicts in isolation. Studies on the intergenerational conflict use a one dimensional voting process where voters decide either on the allocation or the size of government spending on pensions and education (Soares, 2006; Kaganovich and Zilcha, 2012; Naito, 2012). Other studies consider two dimensional voting models where the allocation and the size are determined jointly (Rangel, 2003; Lancia and Russo, 2016; Ono and Uchida, 2016). In the literature on the intragenerational conflict parents are allowed to opt-out of public education by sending their children to private schools, which generates diverging interests between rich and poor (Stiglitz, 1974; Glomm and Ravikumar, 1992; Levy, 2005; de la Croix and Doepke, 2009).

This paper is most related to Naito (2012); Ono and Uchida (2016); Levy (2005) and de la Croix and Doepke (2009). In Naito (2012) these conflicts are boiled down to a political dispute between a coalition of retirees and poor middle-aged and a coalition of rich middle-aged. This study shows that in a repeated majority voting game there is a politico-economic equilibrium where a high initial level of income inequality reduces the size of public education and pensions. Ono and Uchida (2016) consider the intergenerational conflict over pensions and education spending in a probabilistic voting setting. An increase in longevity increases total public pension spending, but the effect of longevity on education is hump shaped. Levy (2005) introduces a model of endogenous political party formation, where there is income redistribution between rich and poor as well as redistribution between young and old in the form of public education. There are four voting groups as agents are differentiated according to their income and age. In this model, if the young are in a minority there is high level of public education provision but the opposite outcome occurs when the young constitute a majority in population. de la Croix and Doepke (2009) show that in an probabilistic voting setting with private and public education, an increase in income inequality that decreases public education participation increases public education quality, but private education can crowd out public education if the political process is dominated by the rich.

We contribute to this literature by augmenting the probabilistic voting model on public and private education developed in de la Croix and Doepke (2009) by the dimension of a pay-as-you-go pension system. This allows us to consider the two political conflicts together and investigate the effect of income inequality and population ageing on education and pension spending. Moreover, we depart from Naito (2012) and Ono and Uchida (2016) by allowing agents to opt-out of public education, and from Levy (2005) by considering pensions for the old. In our model the preferences of heterogeneous agents are aggregated through probabilistic voting. Our goal is to determine simultaneously the size of the government and the allocation of public spending. We find that the education spending per student and pensions per retiree are affected by income inequality and ageing in the same direction. An increase in income inequality increases both per student public education spending as well as public pensions per pensioner, whereas an increase in the share of the population that is retired decreases both public education spending and pensions.

In our overlapping generations (OLG) model agents are heterogeneous with respect to their income. They live for three periods – young, adults (parents) and elderly – and each period they make sequentially two kind of choices, private and public. First, parents decide on the number of children and they choose whether to send them to a public or private school. Afterwards, the electorate (working age adults and pensioners) chooses the level of taxes and their allocation between

Figure 4.1.: *The Life Cycle of Intergenerational Transfers*



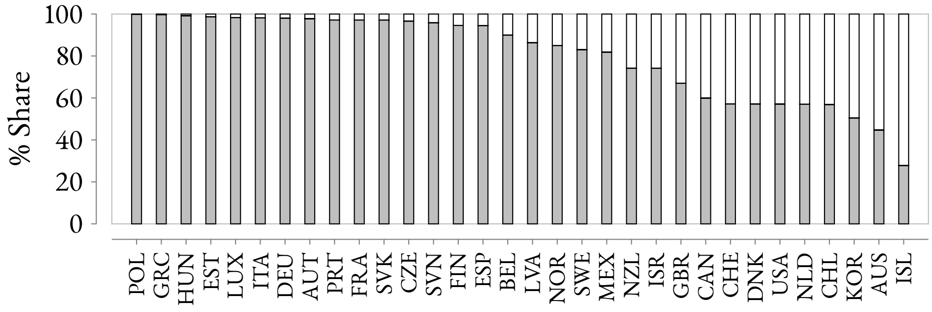
NOTE: This graph depicts the allocation of private and public intergenerational transfers among generations through life. Source: National Transfer Accounts (NTA) data are taken from Ronald Lee and Andrew Mason (2011).

pension and education spending according to a probabilistic voting model (Lindbeck and Weibull, 1987; Persson and Tabellini, 2000). In this setting, on the one hand, an increase in income inequality increases the level of per student public education spending and pensions. On the other hand, an increase in the retired population decreases both the level of public education and pensions. The former operates through the channel of a decreasing public education participation due to the substitution of public by private schooling freeing public resources for higher per student spending. At the same time, some of the resources that are not used for public schooling any more are used in order to finance more generous pensions. The latter works directly via the budget constraint. The increased proportion of elderly burdens the government's budget, inducing cuts in the expenditure on pensions and education per beneficiary.

We conduct a panel data analysis using OECD countries to examine if an increase in income inequality increases, and population ageing decreases public spending per student in primary and secondary education. More specifically, we employ two different specifications, a fixed effects approach and a dynamic panel analysis. We find evidence in favour of a negative effect of population ageing on education spending per student, but we obtain mixed results regarding the effect of income inequality.

Our theoretical approach is motivated by the shape of public and private intergenerational transfers depicted in Figure 4.1. The working age adults pay for the young through both public and private transfers, but for the retired population entirely through public transfers. Figure 4.2 presents further evidence for this: for almost all countries the vast majority of pensions spending is publicly provided.

Figure 4.2.: *Public and Private Pension Spending*



NOTE: Pension spending is defined as all cash expenditures (including lump-sum payments) on old-age and survivors pensions. Source: Pension spending, OECD (2019).

Therefore we choose this particular setting where there is public and private education for the young, but only a public pay-as-you go pensions system for the elderly.[1]

Moreover, as we can see from Figure 4.3 and Figure 4.4, the old dependency ratio (the ratio of retirees that have to be supported by working age adults, henceforth ODR) has increased substantially and it is expected to grow even stronger in the near future.[2] Parallel to the ageing of the population, there was a strong increase in income inequality, leading to an even stronger increase in resources available for education to high income households and a sharp decrease in the resources available to low income households. As can be seen in Figure 4.3, the Gini index as a measure of pre-tax and transfers income inequality has increased for all observed countries. As a result of these trends we expect the intensity of the two political conflicts – intergenerational and intragenerational – over the welfare state to be increasing.

The first political conflict belongs to the literature of the political economy of social security (i.e., public pensions). In this literature, the ageing process affects

[1]In our model, the consumption of the retirees is covered by pensions rather than private savings, which constitute only a fraction of the elderly income in OECD countries (see OECD, 2017).

[2]The main forces behind population ageing are, declining fertility rates after the post-war "baby boom" and increased life expectancy. Among other things, the latter is a result of better quality services due to technological progress in the healthcare system, while the former results from the increasing opportunity cost for women of having children in developed economies. According to Galor and Weil (1996), this is brought about by the higher increase in female wages with respect to household income. Other potential channels include the increase in human capital investment per child and the quantity-quality trade-off à la Becker (1960) (Becker et al., 1990; Galor and Moav, 2000).

pensions through two opposing channels. On the one hand, there is the "fiscal leakage" hypothesis, which suggests that the increasing proportion of elderly decreases the expected profitability of pay-as-you-go pension systems for current working-age voters, thereby inducing them to favour lower current pensions. Therefore, the working-age generation repudiates the social security system (Breyer and Stolte, 2001; Razin et al., 2002; Razin and Sadka, 2007). On the other hand, according to the median voter theorem, governments implement the distribution of public funds that is preferred by the median voter (Downs, 1957) and as the median voter becomes older – due to population ageing – the political clout of the elderly seems set to grow. In turn, the increasing political power of the elderly transforms the allocation of public resources, shifting more resources towards the older cohorts (e.g. for pensions) and fewer to the younger cohorts (e.g. for education) (Browning, 1975). In the context of a limited fiscal budget, this reallocation of public funds might trigger a "struggle" for fiscal resources between the young and elderly, the so-called "intergenerational conflict" hypothesis (Poterba, 1997; Cattaneo and Wolter, 2009; Krieger and Ruhose, 2013).[3]

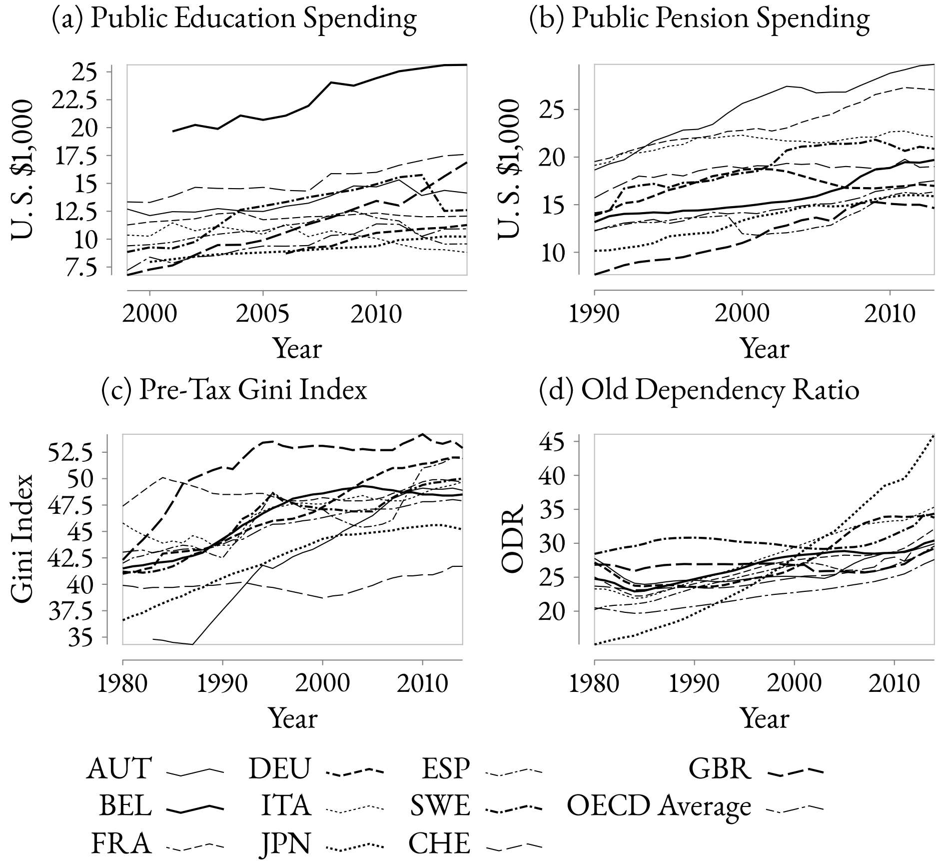
However, it has been pointed out by Casamatta and Batté (2016) that it is crucial to examine the nature of the linkage between publicly funded education and pensions before attempting to predict the effect of ageing on them. Becker and Barro (1988) consider this connection as an exchange of transfers between young and old, where the former pay social security contributions and the latter invest in education. In the same vein Rangel (2003) and Boldrin et al. (2005) consider a type of intergenerational contract in which generations link forward (e.g. education) to backward intergenerational transfers (e.g. pensions) in order to achieve an optimal and sustainable allocation of public economic resources. In particular, Rangel (2003) demonstrates the imperative role of backward intergenerational transfers in sustaining forward intergenerational transfers.[4]

Furthermore, the seminal paper of Pogue and Sgontz (1977) shows that the design of the PAYG pension system – pay contributions "now" and receive benefits "tomorrow" – and consecutively the connection of old age benefits to labour productivity of the future generations – the positive link between pensions and education – generates the appropriate incentives to invest in public education. More

[3] In the literature this hypothesis is also known as the "political power of elderly" (Boadway and Wildasin, 1989; Breyer and Craig, 1997; Tabellini, 2000; Disney, 2007; Shelton, 2008; Tepe and Vanhuysee, 2009).

[4] The political economy application of this theory is empirically evaluated in Michailidis and Patxot (2018).

Figure 4.3.: *Trends in Demographics, Inequality, and Education and Pensions*



NOTE: These plots show the increasing trends in education spending per student and pensions spending per pensioner measured in U. S. $1,000 (PPP 2011), pre-tax and transfers income inequality and old dependency ratio. Data Source: OECD, United Nations and the Standardized World Income Inequality Database. The time span of the graphs is dependent on data availability.

specifically, the working age generations are willing to pay for public education only if they can "reap" gains of higher (human capital) productivity in the future in terms of higher taxable income (Konrad, 1995), social security contributions (Kemnitz, 2000) and/or higher returns on savings (Gradstein and Kaganovich, 2004). Moreover, Lancia and Russo (2016) argue that adults support education only if they can ensure that they will be able to extract a political rent in form of future pensions. Hence, the strategic role of human capital is more important when the political power of the elderly is larger and the forward looking adults support public education policy as they are democratically entitled to claim share of the produced human capital of future generations.[5]

The second political conflict that we are interested in is the intragenerational conflict between rich and poor. Since the 1970s, there was a strong increase in income inequality in the OECD countries (see Piketty, 2013). In the U. S. this has taken the form of a polarisation of incomes (Goos et al., 2009; Acemoglu and Autor, 2011)[6] and parallel to this there was an increase in the inequality of investments into children and the achievement gap between poor and rich students (Kornrich and Furstenberg, 2013; Reardon, 2011).[7] In a similar vein, Mayer (2002) finds that in the U. S. states with higher income inequality have higher differences in educational attainment between children from poor and rich backgrounds, but higher per pupil public education expenditures.[8]

There is a vast literature on income inequality, education and voting. Stiglitz (1974) discusses the effect of different educational institutional arrangements (public v. s. private education) on educational outcomes in a setting with majority voting. He shows that the equilibrium outcome is depending on whether education is mainly understood as a private good or a public good. Bearse et al. (2005) study the effect of income inequality on public and private education in a majority voting model where public education can be both substituted and supplemented by private education expenditures. If supplementary private education spending and

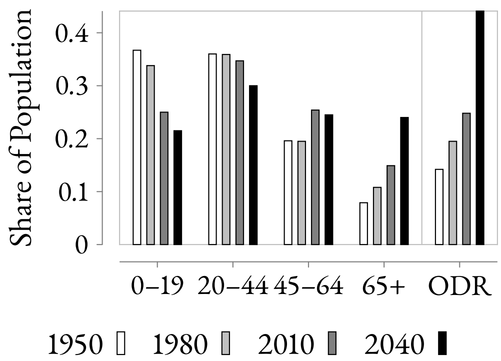
[5] See Michailidis et al. (2019) for the empirical confirmation of this theoretical prediction.

[6] There is no evidence of a polarisation of wages in Europe yet. There is an increase of upper tail inequality, but no decrease of lower tail inequality in the U. K. and Germany (Manning et al., 2007; Antonczyk et al., 2010).

[7] Reardon (2011) shows that parallel to the increase in income inequality in the U. S. there was an increase in the education achievement gap between children from the 90th and the 10th income percentile, though rising income inequality appears not to be the dominant factor.

[8] Bailey and Dynarski (2011) show that there was a strong increase in the college completion rate between 1979 and 1997, with a much stronger increase for children from high income families. This is driven by a strong increase in the college attendance rate of women from high income families.

Figure 4.4.: *Changing Demographic Structure of Voting Cohorts*



Note: The bar plot illustrates the changing demographic structure in OECD countries on average. We divide the total population in 4 major age cohorts: A) Children: Children under 20 years old, B) Young Adults: people from 20 to 44 years old, C) Old Adults: People from 45 to 64 years old, D) Elderly: people above 65 years old. Every age cohort is expressed as share of total population. E) ODR: the share of elderly (over 65 years old) over the working population (20-64 years old). The share of each cohort is depicted over 90 years (1950 to 2040) demonstrating the demographic transition.

private schooling are prefect substitutes, there is no private school enrolment. In a mixed equilibrium, where they are not perfect substitutes, an increase in income inequality first increases per student public education spending, but then decreases it as students start to drop out of private education. Ichino et al. (2011) has a model of social mobility and public education spending. When the poor families are less politically active, there is less public education spending and less social mobility.

Another strand of the literature uses education to link income inequality to economic growth. In Galor and Zeira (1993) and Moav and Galor (2004), credit constraints hinder poor families from acquiring an optimal level of education, which leads to a negative effect of income inequality on economic growth. Other strands of the literature find a negative link between inequality, education and growth through assortive mating (Fernández and Rogerson, 2001) or technological progress (Galor and Tsiddon, 1997). The most related study to us, Glomm and Ravikumar (1992), shows in an endogenous growth model with majority voting that if income inequality is high a public education regime leads to higher growth, whereas if income inequality is low a private education regime leads to higher growth.

The rest of the paper is structured as follows: Section 4.2 introduces our model, Section 4.3 analyses the effect of income inequality and population ageing on the equilibrium levels of public education and pensions, Section 4.4 evaluates these effects using OECD data, and Section 4.5 concludes the paper.

4.2. The Model

Our model based on de la Croix and Doepke (2009) is populated by a continuum of agents that has a mass of one. They live for three periods: in the first period they are born and children, in the second they are adults and work, and in the third they receive a pension and live from that pension. Agents that are working adults in period t base their decisions on the following utility function:

$$\ln(c_t) + \gamma\left[\ln(n_t) + \eta\ln(h_t)\right] + \beta\mathbb{E}\left[U^o_{t+1}(p_{t+1})\right], \tag{4.1}$$

where $U^o_{t+1}(p_{t+1})$ is their utility when old:

$$U^o_{t+1}(p_{t+1}) = \ln(p_{t+1}). \tag{4.2}$$

Here, c_t is the consumption of the agent as adult, p_{t+1} is the pension which they consumes as retiree, n_t is the number of children they have, and h_t is the education of their children in terms of per child education spending. In this model we consider the pension spending per pensioner and education spending per student as the "quality" of pensions and education, respectively. The parents are altruistic towards their children with parameter γ and care about the quality of their children's education relative to the number of children with parameter η. β is the discount factor for the future consumption, and future consumption is equal to the expected pension p_{t+1} that the agent receives.

There are no savings in this economy, and the consumption after retirement is financed through a pay-as-you-go pension system. The agent's budget constraint is equal to

$$c_t + (1 - v_t)n_t e_t = (1 - v_t)y_t(1 - \phi n_t), \tag{4.3}$$

where y_t is the wage, v_t is the income tax rate. ϕ is the per child time that an agent has to dedicate to child rearing, and $1 - \phi$ is the time that an agent works. e_t is the private education spending per child, which is tax exempt, therefore $(1 - v_t)n_t e_t$ is the total private spending on education. We distinguish between agents that send their children to public education, denoted by a superscript s, and agents that send their children to private education, denoted by a superscript e. If parents are sending their children to private education they have to choose the per child spending on education e_t that they have to pay themselves and $h_t = e_t$. If they send their children to public education the level of education is decided and provided for by the government and $h_t = s_t$, where s_t is a political variable. The agents cannot supplement public education by private spending, and $e_t = 0$ for agents with children in public education. The budget constraint for parents sending their children to public education is thus:

$$c_t = (1 - v_t)y_t(1 - \phi n_t).$$

There is no capital in this economy, the potential economic output Y_t (when all agents are employed full time) is equal to a Cobb-Douglas production function using privately and publicly educated agents. The relationship between potential output Y_t and education is defined in the following way:

$$\ln Y_t = \ln A + (1 - \Psi_{t-1}) \ln \hat{e}_{t-1}^{\alpha} + \Psi_{t-1} \ln s_{t-1}^{(1-\alpha)}, \tag{4.4}$$

where $\hat{e}_{t-1}$ is the average spending per student in private education, s_{t-t} is the spending per student in public education, and $\alpha \in [0, 1]$ is the elasticity of substitution between the two. We introduce the share of public education Ψ into the Cobb-Douglas parameter in order to ensure the marginal return on an increase in the spending per student in both the public and the private education sector increases with the number of students attending public and private education respectively. This is needed to guarantee the tractability of the model. This also allows for the existence of a total private education system and a total public education system. A is a parameter that captures the technology and non human capital related parts of the economy. Only adults work, therefore the output depends on the human capital accumulated in the previous period. Individuals differ in the relative share of the total income x that they receive. We normalise the distribution $G(x)$ of x to have mass one, therefore the income that an individual with x could get if they worked full time is equal to

$$y_t = xY_t.$$

We assume that the distribution of x is independent of the choices of last period.

Private and public choices do affect the level of potential income in the future, the relative population size, but not the income distribution. Therefore the distributional parameters stay constant over time, and the political choice in t becomes a static problem independent of the future income distribution and future political choices.

The next period potential output is a function of this period's decisions. In order to solve this model, we assume that the expected value of next periods pensions is proportional to the output of the economy:

$$\mathbb{E}_t(p_{t+1}) \propto Y_{t+1}.$$

That means that if the next periods output increases, agents expect to have an increase in their pensions of the same magnitude as well. This assumption refers to the positive intergenerational link between the working age adults and children.[9] In particular, we assume that it is of the following form:

$$\mathbb{E}_t(p_{t+1}) = \Theta_{t+1} Y_{t+1},$$

where Θ_{t+1} is the expected share of potential output that is dedicated to pensions, a variable that captures the expected future policies. We assume, as standard in the political choice literature, that current policies and decisions do not affect expected future policies, i. e. that Θ_{t+1} is independent of choices made in t.

4.2.1. The Private Choice

Agents optimise their utility over the number of children n_t, their consumption c_t, and the investment into their children's education h_t given their budget constraint (4.3). They take political variables as exogenously given. We distinguish between agents that choose public education for their children, and agents that choose private education for their children, denoted by superscript s and e respectively. If an agent chooses to send their children to public education, they will receive an education in the value of s_t, which will be paid and determined by the government (i. e. the political process). If they send their children to private education, they can choose the level of education spending e_t but have to pay for it themselves.

[9] The working age adults are willing to pay for the education of young because they expect to reap the gains of higher productivity during their retirement in the near future (Konrad, 1995; Kemnitz, 2000)

Incorporating (4.2), (4.3), and (4.4) into utility (4.1), we get the following indirect utilities in the cases of private and public education:

$$\begin{aligned} U_t^s(y_t, n_t | s_t, v_t, p_{t+1}) &= \ln(1 - v_t) + \ln(y_t) + \ln(1 - \phi n_t) + \gamma \ln(n_t) \\ &\quad + \gamma\eta \ln(s_t) + \beta \mathbb{E}\left[\ln(p_{t+1})\right], \qquad (4.5) \\ U_t^e(y_t, n_t, e_t | v_t, p_{t+1}) &= \ln(1 - v_t) + \ln\left[y_t(1 - \phi n_t) - n_t e_t\right] + \gamma \ln(n_t) \\ &\quad + \gamma\eta \ln(e_t) + \beta \mathbb{E}\left[\ln(p_{t+1})\right]. \qquad (4.6) \end{aligned}$$

There is a Beveridgean redistributive pay-as-you-go pension system and agents do not choose the level of pension, which is a political variable. They optimise their utility only over consumption, number of children, and in case they are choosing private education the education spending per child. The optimal choice of variables for parents choosing private education is equal to:

$$\begin{aligned} c_t^e &= (1 - v_t)\frac{y_t}{1 + \gamma}, \\ n^e &= \frac{\gamma(1 - \eta)}{\phi(1 + \gamma)}, \qquad (4.7) \\ e_t^e &= \frac{\eta\phi y_t}{1 - \eta}, \qquad (4.8) \end{aligned}$$

where $n_t^e = n^e$ is static and independent of other variables. The optimal choice for parents choosing public education is equal to:

$$\begin{aligned} c_t^s &= (1 - v_t)\frac{y_t}{1 + \gamma}, \\ n^s &= \frac{\gamma}{\phi(1 + \gamma)}, \qquad (4.9) \end{aligned}$$

where $n_t^s = n^s$ is static and independent of other variables as well.

Agents choose private education if the value of private education in terms of utility is larger or equal to the value of public education in terms of utility, i. e.:

$$U^e(y_t, c_t^e, n^e, e_t^e | v_t, p_{t+1}) \geq U^s(y_t, c_t^s, n^s | s_t, v_t, p_{t+1}). \qquad (4.10)$$

These indirect utilities only depend on y_t, which is directly proportional to x. Agents differ only in the share of total output x that they receive. Thus there will be a $\tilde{x}_t$ for which the utilities in both education systems will be the same. Solving

(4.10) for $\tilde{x}_t$ we get:

$$\tilde{x}_t = \frac{1-\eta}{\hat{\eta}\phi\eta}\mathbb{E}_t(s_t), \tag{4.11}$$

where $\hat{\eta} = (1-\eta)^{1/\eta}$. Here, $\mathbb{E}_t(s_t)$ is the expected value of public education. Agents do not know the realisation of the quality of public education when they decide on fertility and whether they send their children to public or private education. Therefore $\tilde{x}_t$, the x of the agent that is indifferent between sending their children to public or private education depends on the school quanlity that they expect when the agents make their private choice.

We assume a uniform distribution of x over the interval $[1-\sigma, 1+\sigma]$. Therefore the fraction of children participating in the public education system is equal to

$$\Psi_t = \begin{cases} 0 & \text{if} \quad \tilde{x}_t < 1-\sigma, \\ \frac{\tilde{x}_t-(1-\sigma)}{2\sigma} & \text{if} \quad 1-\sigma \leq \tilde{x}_t \leq 1+\sigma, \\ 1 & \text{if} \quad \tilde{x}_t > 1+\sigma. \end{cases} \tag{4.12}$$

In the first case, the x with which an agent would be indifferent between public and private educaiton is lower than the one of the poorest agent in the economy and therefore the share of parents sending their children to public education is equal to 0. In the last case, $\tilde{x}_t$ is larger than the one of the richest agent in the economy, and therefore everyone sends their children to public schools ($\Psi_t = 1$). In the case with $1-\sigma \leq \tilde{x}_t \leq 1+\sigma$ some parents send their children to public and some to private schools.

We define N_t as the population size of the adult at the time t. We define the population growth rate as ϱ_t, such that the relation between population in t and $t-1$ is equal to

$$N_t = (1+\varrho_{t-1})N_{t-1}.$$

We normalise the adult population at t to one, so in t the retired population size of generation $t-1$ is equal to $1/(1+\varrho_{t-1})$. The population growth rate depends on the participation in public education Ψ_t in the following way:

$$1+\varrho_t = \Psi_t n^s + (1-\Psi_t)n^e. \tag{4.13}$$

Since agents that choose public education do not have to pay the cost of education for their children, they choose to have a higher number of children ($n^s > n^e$), and thus an increase in the participation in public education Ψ_t leads to an increase in population growth ϱ_t.

4.2.2. Public Choice

After making their private choices, i. e. deciding whether to participate in public or private education and how many children to have, the adult and the retired agents vote on the public choice variables s_t, p_t, and v_t. A policy $\{s_t, p_t, v_t\}$ has to fulfil the following government budget constraint:

$$\int_0^{\tilde{x}_t} s_t n^s g(x)dx + \frac{1}{1+\varrho_{t-1}} p_t = v_t \left\{ \int_0^{\tilde{x}_t} x(1-\phi n^s) g(x)dx + \int_{\tilde{x}_t}^{\infty} \left[x(1-\phi n^e) - e_t^e(x) n^e \right] g(x)dx \right\}, \quad (4.14)$$

where $g(x)$ is the probability density function of $G(x)$. The left hand side of this equation represents the government expenditures, i. e. the expenditures for public education (first term on the left) and the expenditures for pension of the retired (second term on the left). The right hand side represents the revenue from income taxes v_t on those with public education (first term on the right) and those with private education (second term on the right). Using (4.7), (4.8), and (4.9) we can show that the taxable income in period t is equal to

$$\int_0^{\tilde{x}_t} x\left(1-\phi n^s\right) g(x)dx + \int_{\tilde{x}_t}^{\infty} \left[x\left(1-\phi n^e\right) - e(x) n^e \right] g(x)dx = \frac{Y_t}{1+\gamma} \int_0^{\infty} x g(x)dx = \frac{Y_t}{1+\gamma}. \quad (4.15)$$

where $e(x) = e_t^e$ for agents with income $y_t = xY_t$. The tax revenue is independent of the participation rate Ψ_t and only depends on the economic output. Using this, we can rewrite the government budget constraint (4.14) as

$$v_t \frac{Y_t}{1+\gamma} = s_t \Psi_t n^s + p_t \frac{1}{1+\varrho_{t-1}},$$

which leads to the following expression of the tax rate v_t as a function of per pensioneer pensions p_t and per student spending on public education s_t

$$v_t = \frac{1+\gamma}{Y_t}\left(s_t \Psi_t n^s + p_t \frac{1}{1+\varrho_{t-1}}\right). \tag{4.16}$$

Thus we can replace v_t in the indirect utilities (4.5) and (4.6) with (4.16) and formulate the public decision as a decision on two variables p_t and s_t, where the tax rate v_t is a function of the two. The policy variables are chosen according to a probabilistic voting, where the adults and retirees vote on competing political platforms defined on $\{s_t, p_t\}$ (for a discussion of the probabilistic voting see Appendix C.1). The winning political platform is the one that optimises the following objective function:

$$\Omega(s_t, p_t) = \int_0^{\tilde{x}_t} U_t^s \left[x, s_t, p_t, v_t(s_t, p_t)\right] g(x)dx + \int_{\tilde{x}_t}^{\infty} U_t^e \left[x, s_t, p_t, v_t(s_t, p_t)\right] g(x)dx + \frac{1}{1+\varrho_{t-1}} U_t^o(p_t).$$

One can show that Ω is strictly concave in s_t and p_t. The maximisation of Ω with respect to s_t leads to

$$0 = -\frac{\Psi_t n^s}{\frac{Y_t}{1+\gamma} - s_t \Psi_t n^s - p_t \frac{1}{1+\varrho_{t-1}}} + \Psi_t \frac{\eta\gamma}{s_t} + \frac{\beta \Psi_t (1-\alpha)}{s_t}. \tag{4.17}$$

The first term on the right is the costs of an increase in s_t through taxes for the adult population, the second term is the benefit of an increase in s_t for the parents sending their children to public schools, and the third term is the benefit of an increase s_t for all adults through the higher expected future production that is paying for their pensions.

Maximising Ω with respect to p_t yields

$$0 = -\frac{\frac{1}{1+\varrho_{t-1}}}{\frac{Y_t}{1+\gamma} - s_t \Psi_t n^s - p_t \frac{1}{1+\varrho_{t-1}}} + \frac{1}{1+\varrho_{t-1}} \frac{1}{p_t}. \tag{4.18}$$

Again, the first part of this equation represents the costs of an increase in p_t through taxes on adults income and the second part the benefit of an increase in p_t for the retirees.

We can now use (4.17) and (4.18) to solve for the political outcome of the voting process $\{s_t^*, p_t^*\}$:

$$s_t^* = \frac{(1+\varrho_{t-1})\left[\eta\gamma + \beta(1-\alpha)\right]}{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma + \beta(1-\alpha)+1\right]+1}\frac{Y_t\gamma}{\phi}, \tag{4.19}$$

$$p_t^* = \frac{1}{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma + \beta(1-\alpha)+1\right]+1}\frac{Y_t(1+\varrho_{t-1})}{1+\gamma}. \tag{4.20}$$

We can insert (4.19) and (4.20) into (4.16) to get the tax rate v_t^* that corresponds to this policy:

$$v_t^* = \frac{\Psi_t\gamma\eta + \frac{1}{1+\varrho} + \Psi_t\beta(1-\alpha)}{1 + \Psi_t\gamma\eta + \frac{1}{1+\varrho} + \Psi_t\beta(1-\alpha)}. \tag{4.21}$$

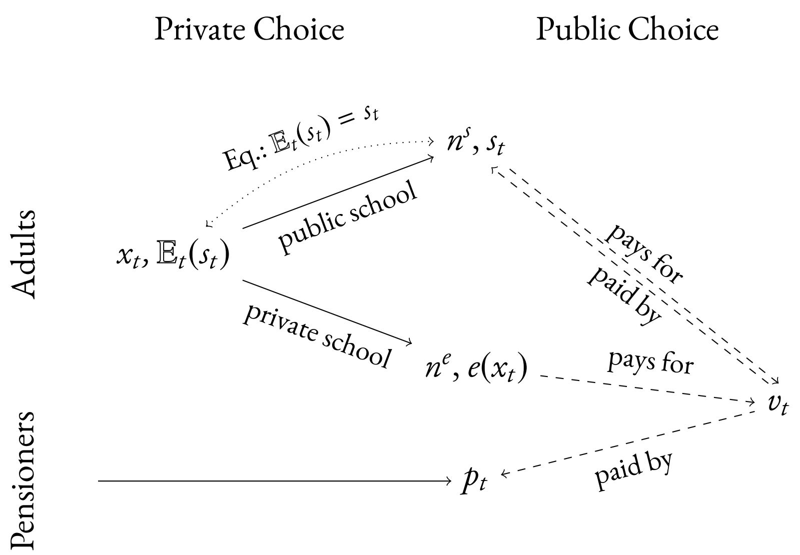
According to the probabilistic voting theory, it is optimal for competing political platforms to offer the policy $\{s_t^*, p_t^*, v_t^*\}$, which is maximising the probability of being elected. Therefore this is the equilibrium outcome of the political process. All these political variables are dependent on the participation rate in public education Ψ_t, which is an outcome of the expectations on the level of public schooling $\mathbb{E}_t(s_t)$. A representation of this sequence of the above choices is depicted in Figure 4.5. We are now going to define an equilibrium with perfect foresight of the agents with respect to s_t.

4.2.3. Equilibrium

In this model, agents are deciding first whether or not to send their children to public education based on their expectations on the level of public education ($\mathbb{E}(s_t)$). This decision then influences the outcome of the political process and thus the level of public education s_t itself. We are assuming perfect foresight of the agents with respect to this periods policies, and an equilibrium is thus defined as the expected value of s_t that yields itself as the outcome of aggregated private choices and the resulting public policies:

Definition 1. *An equilibrium consist of an income threshold $\tilde{x}$ satisfying (4.11), a fertility rule $n = n^s$ for $x \leq \tilde{x}$ and $n = n^e$ for $x > \tilde{x}$, a private education decision $e = 0$ for $x \leq \tilde{x}$ and $e = e^e(x)$ for $x > \tilde{x}$, and aggregate variables $\{\Psi_t, s_t^*, p_t^*, v_t^*\}$ given by equations (4.12), (4.19), (4.20), and (4.21), such that the perfect foresight*

Figure 4.5.: *Sequence of Choices*



Note: First, adults choose whether to send their children into public or private schools and how many children to have (n^s or n^e), as well as the level of private education $e(x_t)$ in case their children attend a private school. This private decision depends on their location in the income distribution x_t and the expected per student spending in public schools $\mathbb{E}(s_t)$. Afterwards the electoral body (adults and pensioners) vote simultaneously on the tax rate v_t, per pensioner pensions p_t and per student spending in public schools s_t. An equilibrium of this model is the point where the expectations are fulfilled, i. e. $\mathbb{E}(s_t) = s_t$.

condition holds:

$$\mathbb{E}_t(s_t) = s_t. \tag{4.22}$$

To show that an equilibrium exists and is unique, we are using Brouwer's fixed-point theorem. For this we need the following lemma:

Lemma 2. *The level of public education s_t^* and the level of public pensions p_t^* are decreasing in the participation in public education Ψ_t, whereas the tax rate v_t^* is increasing in participation in public education.*

Proof. The first derivative of s_t^* and p_t^* with respect to Ψ_t are equal to

$$\frac{\partial s_t^*}{\partial \Psi_t} = -\frac{(1+\varrho_{t-1})^2\left[\eta\gamma + \beta(1-\alpha)\right]\left[\eta\gamma + \beta(1-\alpha)+1\right]}{\left\{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma + \beta(1-\alpha)+1\right]+1\right\}^2}\frac{Y_t\gamma}{\phi}, \tag{4.23}$$

and

$$\frac{\partial p_t^*}{\partial \Psi_t} = -\frac{(1+\varrho_{t-1})^2\left[\eta\gamma + \beta(1-\alpha)+1\right]}{\left\{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma + \beta(1-\alpha)+1\right]+1\right\}}\frac{Y_t}{1+\gamma}, \tag{4.24}$$

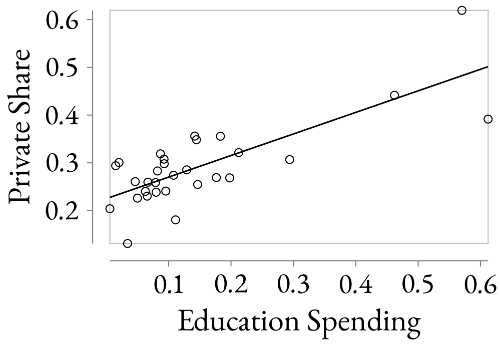
which are both always negative. The first derivative of v_t^* with respect to Ψ_t is equal to

$$\frac{\partial v_t^*}{\partial \Psi_t} = \frac{\gamma\eta + \beta(1-\alpha)}{\left[1 + \Psi_t\gamma\eta + \frac{1}{1+\varrho} + \Psi_t\beta(1-\alpha)\right]^2}, \tag{4.25}$$

which is always positive. □

A decrease in the participation in public education Ψ_t means that there are now less parents that are voting in favour of public education, and also the weight of public educated children in the future production is decreasing. But at the same time the number of children in public education is decreasing, which is dominating the other effect here. Since with the decrease in the number of children a higher level of public education can be provided for a lower costs, there are more funds to increase the level of pensions and decrease the tax rate. This is in line with empirical evidence for OECD countries as shown in Figure 4.6, there is a positive correlation of 0.77 (0.000) between participation in private education and per student spend-

Figure 4.6.: *Participation and Per Student Spending in Public Education*



NOTE: This scatter plot depicts the relationship between the private share in primary & secondary education and public education spending per student in primary & secondary education as a share in GDP per capita, for all countries in our sample. This relationship is highly correlated and statistically significant 0.77 (0.000).

ing in public education.[10]

Now, we are using Lemma 2 to show that an equilibrium exists and is unique.

Proposition 10. *An equilibrium exists and is unique.*

Proof. The existence and uniqueness of an equilibrium as defined in Definition 1 follow from an application of the Brouwer's fixed-point theorem. Using (4.19), the actual quality s_t and the expected schooling quality $\mathbb{E}_t(s_t)$ lie in the interval

$$\mathbb{E}_t(s_t), s_t \in \left\{ \frac{(1+\varrho_{t-1})\left[\eta\gamma + \beta(1-\alpha)\right]}{(1+\varrho_{t-1})+1} \frac{Y_t\gamma}{\phi}, \right. \\ \left. \frac{(1+\varrho_{t-1})\left[\eta\gamma + \beta(1-\alpha)\right]}{(1+\varrho_{t-1})\left[\eta\gamma + \beta(1-\alpha)+1\right]+1} \frac{Y_t\gamma}{\phi} \right\}. \tag{4.26}$$

We define a mapping Δ from $\mathbb{E}_t(s_t)$ into s_t, which maps this interval into itself. A unique fixed point of this mapping implies the existence of a unique equilibrium with $\mathbb{E}_t(s_t) = s_t$. Using (4.11) and (4.12), we can show that the participation in public education $\Psi - t$ as a function of $\mathbb{E}_t(s_t)$ is equal to:

$$\Psi_t = \Psi[\mathbb{E}_t(s_t)] = \max\left\{ \min\left[\frac{1-\eta}{2\sigma\hat{\eta}\phi\eta}\mathbb{E}_t(s_t) - \frac{1-\sigma}{2\sigma}, 1 \right], 0 \right\}. \tag{4.27}$$

This function is weakly increasing in $\mathbb{E}_t(s_t)$. The higher the expected quality of public education, the more parents are going to prefer sending their children to

[10] de la Croix and Doepke (2009) find this as well for the U. S. regions.

public education.

We can use (4.19) to define the mapping Δ, which gives us the actual per student public education expenditure s_t that results for the voting process with the participation rate $\Psi[\mathbb{E}_t(s_t)]$ from (4.27). This education quality $s_t = \Delta[\mathbb{E}_t(s_t)]$ is given by

$$\Delta[\mathbb{E}_t(s_t)] = \frac{(1+\varrho_{t-1})\left[\eta\gamma + \beta(1-\alpha)\right]}{(1+\varrho_{t-1})\Psi[\mathbb{E}_t(s_t)]\left[\eta\gamma + \beta(1-\alpha)+1\right]+1}\frac{Y_t\gamma}{\phi}. \tag{4.28}$$

An equilibrium is a fixed point of $\Delta[\mathbb{E}_t(s_t)]$, i. e. public education spending s_t that satisfies $s_t = \Delta(s_t)$. At this fixed point the schooling quality s_t that is expected by the agents is identical to the one that results from the voting process. Given (4.28) and Lemma 2, Δ is a continuous, weakly decreasing function mapping the closed interval given in (4.26) into itself. The mapping therefore crosses the 45 degree-line exactly once, and a unique equilibrium exists. □

This proof of the existence and uniqueness of the equilibrium works in the following way: according to Lemma 2 the equilibrium per student spending on public education is decreasing with the participation rate in public education. As $\Psi_t \in [0,1]$, the level of the per student spending on public education s_t^* is also bounded. Because the participation rate is an increasing function of the expected schooling quality, and the actual schooling quality is a decreasing function of the participation rate in public education, the actual schooling quality is a decreasing function of the expected schooling quality. As the actual schooling quality is decreasing in expected schooling quality, and both are bounded, according to Brouwer's fixed-point theorem there exists a unique fix point between the two. This is the equilibrium point where expected schooling quality and actual schooling quality coincide and the perfect foresight condition holds.

4.3. Comparative Statics

We can now use the equilibrium schooling and pensions to derive comparative statics in the model. In particular, we are interested in the effect of changes in income inequality on public education provision and pensions. There are three different education regimes: (i.) majority public with $\Psi_t \in [1, 1/2)$; (ii.) equally separated with $\Psi_t = 1/2$; or (iii.) majority private with $\Psi_t \in (1/2, 0]$. Unlike de la Croix and Doepke (2009) we cannot rule out any of this regimes, but as can be seen in Figure 4.7 almost all countries have majority public education regimes, and therefore we concentrate our analysis on this case (for an analysis of the other regimes see Appendix C.2). Initially we are looking at the effect of income inequality on the participation rate in public education. We get for the relationship between the inequality σ and Ψ_t the following:

Proposition 11. *In a majority public education regime with* $\Psi_t > 1/2$ *participation in public education* Ψ_t *and the tax rate* v_t^* *are decreasing with income inequality* σ *and the quality of public education* s_t^* *and the pensions per pensioner* p_t^* *are increasing in* σ.

Proof. The first derivative of Ψ_t with respect to σ is

$$\frac{\partial \Psi_t}{\partial \sigma} = \frac{\sigma - \left[\frac{1-\eta}{\hat{\eta}\phi\eta}\mathbb{E}_t(s_t) - (1-\sigma)\right]}{2\sigma^2} = \frac{1}{\sigma}\left(\frac{1}{2} - \Psi_t\right). \tag{4.29}$$

This is negative for $\Psi_t > 1/2$. Following Lemma 2 this means that p_t^* and s_t^* are increasing in σ and v_t^* is decreasing in σ for $\Psi_t > 1/2$. □

The mechanism of the effect of an increase in income inequality is the following: an increase in income inequality is increasing the income of the marginal agent that is indifferent between private and public education if this agent has an above average income. This means that this agent now prefers private education. This decrease in public education perticipation decreases the share of voters with children in public education, but it also decreases the number of children in public education. Therefore the total spending on public education decreases, but the number of children in publice education decreases stronger. Overall this leads to an increase in per student public education spending. The decrease in total education spending leads to an increase in pensions and to a decrease in taxes.

Secondly, we look at the effect of an increase in the share of old people in the population $1/(1+\varrho_{t-1})$ on pensions and per student public education spending. For this we look at the comparative statics of $1/(1+\varrho_{t-1})$ on p_t^*, s_t^*, v_t^*, Ψ_t, and $(1+\varrho_t)$:

Proposition 12. *An increase in the share of retirees in the population* $1/(1+\varrho_{t-1})$ *decreases the pensions per pensioner* p_t^**, the level of public schooling* s_t^**, and the participation in public education* Ψ_t*, and it increases the tax rate* v_t^**. It also decreases future population growth* $(1+\varrho_t)$.

Proof. Using the implicit function theorem, (4.11), (4.12), and (4.23), we can derive the first derivative of s_t^* with respect to $1/(1+\varrho_{t-1})$:

$$\frac{\partial s_t^*}{\partial \frac{1}{1+\varrho_{t-1}}} = -\frac{1}{\frac{1-\eta}{2\sigma\hat{\eta}\phi\eta}[\eta\gamma+\beta(1-\alpha)+1] + \frac{\left\{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma+\beta(1-\alpha)+1\right]+1\right\}^2}{(1+\varrho_{t-1})^2\left[\eta\gamma+\beta(1-\alpha)\right]}\frac{\phi}{Y_t\gamma}}, \tag{4.30}$$

which is always negative. Following (4.11) and (4.12) this leads to a decrease in the equilibrium value of Ψ_t and according to (4.13) this decreases $(1+\varrho_t)$.

Using this, (4.11), (4.12), and (4.24), we can derive first derivative of p_t^* with respect to $1/(1+\varrho_{t-1})$:

$$\frac{\partial p_t^*}{\partial \frac{1}{1+\varrho_{t-1}}} = -\frac{\frac{1}{\eta\gamma+\beta(1-\alpha)}\frac{\phi}{\gamma(1+\gamma)}}{\frac{1-\eta}{2\sigma\hat{\eta}\phi\eta}[\eta\gamma+\beta(1-\alpha)+1] + \frac{\left\{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma+\beta(1-\alpha)+1\right]+1\right\}^2}{(1+\varrho_{t-1})^2\left[\eta\gamma+\beta(1-\alpha)\right]}\frac{\phi}{Y_t\gamma}},$$

which is also always negative.

Following from (4.30), (4.11), (4.12), and (4.24) the first derivative of v_t^* with respect to $1/(1+\varrho_{t-1})$ is

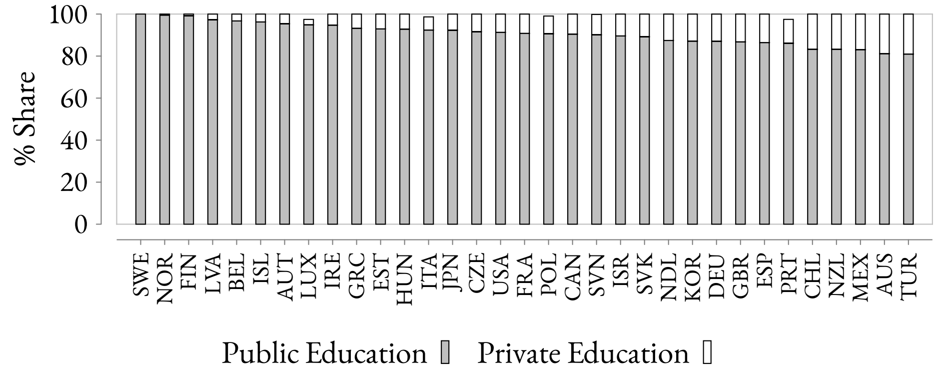
$$\frac{\partial v_t^*}{\partial \frac{1}{1+\varrho_{t-1}}} = \frac{1}{\left[1+\Psi_t\gamma\eta+\frac{1}{1+\varrho}+\Psi_t\beta(1-\alpha)\right]^2} \cdot \frac{\frac{1-\eta}{2\sigma\hat{\eta}\phi\eta} + \frac{\left\{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma+\beta(1-\alpha)+1\right]+1\right\}^2}{(1+\varrho_{t-1})^2\left[\eta\gamma+\beta(1-\alpha)\right]}\frac{\phi}{Y_t\gamma}}{\frac{1-\eta}{2\sigma\hat{\eta}\phi\eta}[\eta\gamma+\beta(1-\alpha)+1] + \frac{\left\{(1+\varrho_{t-1})\Psi_t\left[\eta\gamma+\beta(1-\alpha)+1\right]+1\right\}^2}{(1+\varrho_{t-1})^2\left[\eta\gamma+\beta(1-\alpha)\right]}\frac{\phi}{Y_t\gamma}}.$$

This is always positive. □

The mechanism behind this is similar to the one in Proposition 11: an increase in the share of old people increases the share of voters voting for pensions, but also increases the number of pensioners. This increases the total spending on pensions, but decreases the pensions per pensioner. The increase in pensions is paid by an increase in taxes and a decrease in public education spending. The decrease in public education spending leads to a decrease in participation in public education, which leads to a decrease in population growth.

To conclude the theoretical predictions of the model, an increase in income inequality decreases taxes, but increases per student spending on public education and per pensioner pensions. It decreases the size of the welfare state but increases

Figure 4.7.: *Distribution of Public and Private Funds*



NOTE: Distribution of public and private funds for primary, secondary and post-secondary non-tertiary educational institutions. Final funds after transfers between public and private sectors, excluding international funds (2015). Source: Education at a Glance, OECD, 2018.

the quality of the provided services. On the other hand, an increase in the population weight of the retirees does decrease both the per pensioner pensions and the public education spending per student. Both mechanism operate mainly through fiscal leakage in the budget constraint. An increase in income inequality increases the income of the agent indifferent between public and private education, and thus decreases the participation in public education. This reduces the share of voters caring for public education through altruism for their children, which reduces the total public education spending (which in turn decreases taxes and increases pensions). The number of children attending public education decreases faster than the total spending, which leads to an increase in per student spending on public education. The mechanism in the case of an increase in the number of retirees works in a similar fashion: The increase in the number of pensioners increases the political weight of the retirees, increasing total pension spending (which increases taxes and decreases per student public education spending). The number of pensioners increases faster than the total pension spending, thus the per pensioner pension is decreasing. In both cases we find a positive relationship between per student public education spending and pensions through the budget constraint.

4.4. Empirical Evidence

The theoretical model that we develop in this paper makes prediction on how public education spending per student is affected by income inequality and population ageing. The main predictions of our model about the intergenerational and the intragenerational conflict are the following ones: (i.) Education spending per student and pensions spending per retiree are positively related and affected by changes in inequality and ageing towards the same direction. (ii.) When the majority of children attend public education, a rise in income inequality decreases the participation in public schooling (primary & secondary) and increases the per student spending on education. (iii.) An increase in the share of elderly decreases the per student education expenditures and the per pensioner pensions. We test these theoretical predictions using data on OECD countries in order to assess the validity of our model. The main goal is to investigate how primary and secondary public education spending per student are affected by changes in population ageing and income inequality.

4.4.1. Data

We consider a cross-country analysis using panel data on OECD countries and yearly observations over the period 1998–2014.[11] [12] More specifically, we use aggregated data on public education spending, participation in public and private schooling, income inequality, population ageing and pensions, taken from OECD, UNESCO and World Bank datasets.[13]

As a dependent variable we set the public education spending per enrolled student in only primary public education (henceforth, ESPSPE), only secondary public education (henceforth, ESPSSE), as well as the total primary and secondary public education spending (henceforth, ESPSPSE). Education expenditure is calculated by dividing the total general government expenditure on only primary, only secondary, and total expenditure on primary and secondary education – measured in $ PPP (constant 2011) – by the number of the enrolled students in only primary public education, only secondary public education, as well as the total enrolments in public primary and secondary education, respectively. We also use

[11]OECD countries in our sample: Australia, Austria, Belgium, Chile, Canada, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Latvia, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, South Korea, Spain, Sweden, Switzerland, Turkey, the U.K. and the U.S.. We exclude from our OECD sample Canada and the newest OECD member Lithuania, due to the missing data.

[12]As it is pointed out by (de la Croix and Doepke, 2009) it is a common sense to assume that governments adjust their budget for education on a yearly base.

[13]More detailed description of variables and data sources are provided in the Appendix C.4 Table C.3.

as dependent variable the total government education spending as % of GDP on primary (GEPE), secondary (GESE), and the sum of primary and secondary education (GEPSE). The main results hold for this specification. For the analysis on total education spending as % of GDP, see Appendix C.3 Table C.1.

As main explanatory variables we use the old dependency ratio (ODR) that measures the size of the elderly (population above 65 years old) relative to the size of the working age population (20–64 years old) in order to capture the effect of population ageing on education spending.[14] We use the Gini index (henceforth, Gini) as a measurement of the market income inequality before taxes and transfers to capture the impact of income inequality on education spending. Following de la Croix and Doepke (2009) the Gini coefficient is used in its lagged form in order to avoid possible reverse causality from education to income inequality. More specifically, we use levels of Gini index with a 24 year lag, i. e. the 1975 to 1991 time period of 17 years that correspond to our sample span (1998–2014).[15]

Furthermore, we control for the share of private enrolments – the indirect effect of income inequality on education spending in our model – in only primary, only secondary, and total primary and secondary schooling. Our model predicts that an increase in private schooling participation translates into less students attending public schools and hence higher per student public education spending. Moreover, since public education and pensions compete for the same fiscal resources (intergenerational conflict) we control for the level of pensions. More specifically, we control for pension "generosity" using the level of public pensions per retiree (henceforth, PubPen) as a proxy. Pensions per pensioner are calculated using the total public pensions in % of GDP divided by the number of the people that are expected to be retired (population above 65 years old). Finally, we control for the level of economic development using GDP per capita measured in $ PPP (constant 2011). Table 4.1 displays the descriptive statistics of all variables used in our empirical analysis.

[14] As robustness check we also use a broader measure of old dependency ratio, that is population over 55 years old as a percentage of working age people from 20 to 54 years old. The quantitative results do not change, see Appendix C.4 Table C.4.

[15] We use a 24 year lag following the definition of the UN of "young people" for youth unemployment to ensure that the inequality is measured before the birth of anyone who is still in education.

Table 4.1.: *Descriptive Statistics*

(a) Dependent Variables	N	Mean	St. Dev.	Min	Max
GEPE: Government Expenditure in Primary Education (as % of GDP)	475	1.4045	0.4343	0.5369	2.6773
GESE: Government Expenditure in Secondary Education (as % of GDP)	487	2.0414	0.4414	0.9650	3.0541
GEPSE: Government Expenditure in Primary & Secondary Education (as % of GDP)	472	3.4664	0.6456	2.2461	5.2068
ESPSPE: Education Spending per Student in Primary Education (in $1,000 PPP, constant 2011)	444	8.5155	4.1335	1.6243	27.3467
ESPSSE: Education Spending per Student in Secondary Education (in $1,000 PPP, constant 2011)	440	10.8157	5.5995	2.1625	30.1209
ESPSPSE: Education Spending per Student in Primary & Secondary Education ($1,000 PPP, constant 2011)	420	9.6731	4.6894	1.8134	25.6298

(b) Main Explanatory Variables	N	Mean	St. Dev.	Min	Max
Gini: Gini index pre-tax and transfers (%)	595	47.1395	4.9288	30.8	60.3
ODR: Old Dependency Ratio (Over 65/20-64) (%)	595	24.2664	5.6864	9.9357	46.0558
ODR(20-54): Old Dependency Ratio (Over 55/20-54) (%)	595	30.4627	7.0934	12.0325	52.8460

(c) Control Variables: Public & Private Enrolments	N	Mean	St. Dev.	Min	Max
ENPUBPE: Enrolments in Public Primary Education (in millions)	532	2.3827	4.4629	0.02857	22.5571
ENPUBSE: Enrolments in Public Secondary Education (in millions)	510	2.4966	4.2807	0.0268	22.5634
ENPUBPSE: Enrolments in Public Primary and Secondary Education (in millions)	503	4.9429	8.7864	0.0561	44.8700
SHPRPE: Share of Private Primary Education	515	0.0960	0.1315	0.0008	0.6151
SHPRSE: Share of Private Secondary Education	495	0.1424	0.1390	0.0032	0.6949
SHPRPSE: Share of Private Primary and Secondary Education	486	0.1198	0.1287	0.0055	0.6122

(d) Other Control Variables	N	Mean	St. Dev.	Min	Max
PubPen: Public Pensions per retiree (in $1,000 PPP, constant 2011)	560	15.3260	7.1656	1.5390	44.1942
GDPpc: GDP per capita (in $1,000 PPP, constant 2011)	595	34.7077	14.4755	10.1492	97.8642

Note: Definitions and sources of the data can be found in the Appendix C.4 Table C.3.

4.4.2. Two-way Fixed Effects Model

The cross-country analysis over time (panel analysis) seems to be the most appropriate way to examine empirically the effects of income inequality and population ageing on public education expenditure for primary and secondary education levels. Since income inequality, population ageing, and education spending vary over time and across countries, the standard two–way fixed effects approach fits our purpose. More specifically, the fixed effects assumption is needed in order to avoid systematic biases connected to unobserved characteristics (like culture heritage or religion) that remain constant over years and might have a significant influence on public education spending.[16] The Hausman test points to the use of fixed effects and is in line with our theoretical reasoning.[17] Additional diagnostic tests reveal a need to use time fixed effects and heteroscedastically robust standard errors.[18]

As baseline estimations we use the following two-way fixed effects specification:

$$\ln(Y_{i,t}) = b + \beta X'_{i,t} + \alpha_i + \gamma_t + \varepsilon_{i,t},$$

where $Y_{i,t}$ is public education spending per student of country i at time t, b is the constant term, β is a coefficient vector, and α_i and γ_t represent country and time fixed effects, respectively. Finally, $\varepsilon_{i,t}$ is the idiosyncratic error term. The vector X includes all the regressors used in our estimations.

Table 4.2 shows estimations of the above specified model when we apply the within regression estimator. In the first three regressions we use as dependent variable the log of education spending per student for total (primary and secondary), only primary and only secondary, respectively. Moreover, as main explanatory variable we employ the current (non-lagged) Gini index. In regressions 4 to 6 we use instead the lag of Gini. Regression 1 shows a weak negative effect of current income inequality on public education spending per student for primary and secondary

[16] Castles (1994) argues that cultural heritage and the tradition of Catholicism can play an important role in public expenditure on education. Countries that have Catholicism as their predominant religion might have to spend less on public education of children as the Catholic Church undertakes a large part of the children's education.

[17] More specifically, we reject the the null hypothesis that random effects provide consistent estimates or that there is no correlation between the error term and the independent variables (Hausman, 1978).

[18] We use the time fixed effects test "testparm" available in STATA 14. We reject the null hypothesis: *no time fixed effects*. Also, we conduct the modified Wald test for groupwise heteroskedasticity in the residuals of fixed effects regression introduced by Baum (2001). Again, the null hypothesis: *presence of homoskedasticity*, is rejected.

education when they are considered together. Regression 2 reveals that this negative effect is mainly driven by public primary education spending, as the same effect is insignificant for the secondary education. However, as have mentioned above, the use of the current income inequality may generate problems of reverse causality – from education to income inequality – that we avoid by using a 24 lag of the Gini. When we address this problem – in regressions 4, 5 and 6 – the coefficients of income inequality become positive, although this effect is not significant for only primary education. This result is in line with our theoretical prediction that income inequality decreases the participation in public education increasing the spending per student in public schools.

Our estimations also show that while public pensions per pensioner have the expected positive effect, they are statistically insignificant for estimations with lagged inequality. Moreover, the share of private enrolments has a positive impact on primary and secondary education, but the effect is only significant for the latter. Additionally, the GDP per capita has the expected positive effect on education spending, reflecting the fact that richer countries have higher education spending. Except of the old dependency ratio, the rest of the variables in our estimations behave in the expected way.

As we can see from Table 4.2 the coefficient of the old dependency ratio is positive but is not significant (regressions 4 to 6). However, the effect of old dependency ratio might dependent on the level of pensions per pensioner which could lead to a misspecification of the model.[19] The intuition for this comes directly from the literature on intergenerational conflict where elderly try to appropriate more resources in their favour when there is a competition for fiscal resources. Hence, we estimate our model including the interaction between pensions and old dependency ratio.

Additionally, further diagnostic tests reveal the presence of cross-sectional dependence and autocorrelation in error terms.[20] As mentioned in Cameron and Trivedi (2010), ignoring cross-sectional dependence and correlation of errors over time can lead to systematic bias and thus to erroneous results. To cope with autocorrelation and cross-sectional dependence in the idiosycratic errors we use an

[19]As shown in Appendix C.4 Table C.2, the level of pensions and the old dependency ratio are positively correlated.

[20]More specifically, using Pesaran's cross-dependence test introduced by Pesaran (2004), we reject the null hypothesis: *residuals across entities are not correlated*. Also, using the serial correlation test or the test for autocorrelation by Wooldridge (2010), we reject the null hypothesis: *no serial correlation*.

Table 4.2.: *Ageing and Inequality Effect on Education Spending per Student*

	(1) ESPSPSE	(2) ESPSPE	(3) ESPSSE	(4) ESPSPSE	(5) ESPSPE	(6) ESPSSE
Gini	-0.0149†	-0.0225*	-0.0131			
	(0.008)	(0.010)	(0.009)			
L.24.Gini				0.0186†	0.0216	0.0189*
				(0.011)	(0.017)	(0.007)
ODR	0.0073	0.0000	0.0147*	0.0092	0.0021	0.0145
	(0.008)	(0.010)	(0.007)	(0.010)	(0.014)	(0.009)
PubPen	0.0201*	0.0222**	0.0188*	0.0137	0.0136	0.0188
	(0.007)	(0.007)	(0.009)	(0.011)	(0.008)	(0.012)
GDPpc	0.0588***	0.0638***	0.0510***	0.0543***	0.0500***	0.0541***
	(0.007)	(0.011)	(0.008)	(0.009)	(0.013)	(0.007)
SHPRPSE	1.0795*			0.9953*		
	(0.413)			(0.395)		
SHPRPE		1.5482†			1.5024	
		(0.880)			(0.985)	
SHPRSE			1.1061**			1.0487**
			(0.340)			(0.292)
Ctry. & Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	371	396	391	294	315	304
Countries	32	33	34	31	32	33
F-test	79.90***	21.41***	55.99***	137.37***	65.55***	127.78***
R^2-within	0.8142	0.7829	0.7316	0.7674	0.7264	0.7378

NOTE: Two-way fixed effects regressions with robust standard errors reported in parentheses, ***p<0.001, **p<0.01, *p<0.05, †p<0.10. The standard errors are clustered over the number of countries used in each regression. Dependent variables: education spending per student in primary (ESPSPE), secondary (ESPSSE), primary & secondary education (ESPSPSE) are in logs. Gini: current Gini index on pre tax and transfers income and L.24.Gini is a lag (24 years) of the Gini index, ODR: old dependency ratio. Public pensions spending per pensioner (PubPen) and GDPpc are measured in $1,000 PPP (constant 2011). Share of private education in total primary (SHPRPE), secondary (SHPRSE), primary & secondary (SHPRPSE) education, Constant is not reported but included in all the regressions above.

estimation method that allows us to conduct consistent estimations in the presence of AR(1) autocorrelation within panels and contemporaneous correlation. For that purpose, we use the estimator (SCC) introduced by Hoechle (2007), that produces Driscoll and Kraay (1998) standard errors for the estimated coefficients using fixed effects. In our specification of this estimator, the error structure is assumed to be heteroscedastic, autocorrelated up to one lag and correlated between the countries. As mentioned in Hoechle (2007), Driscoll-Kraay standard errors are robust to very general forms of cross-sectional and temporal dependence when the time dimension is large enough. Additionally, their particular technique to estimate standard errors does not impose any restrictions on the number of countries, which can be even bigger than the number of periods. Moreover, as Cameron and Trivedi (2010) show, the implementation of Driscoll and Kraay's covariance estimator works for both balanced and unbalanced panels. All the above properties make this estimator suitable for our panel data analysis.

In Table 4.3 we make the following changes compared to Table 4.2: First, we introduce the interaction term between old dependency ratio and public pensions per pensioner in order to capture the plausible dependence of the former on the latter in its impact on education spending per student. More specifically, we estimate the first 3 regressions using time fixed effect just as in Table 4.2. Second, we use the estimation technique described above in order to avoid the biased estimates to estimate the same model in regressions 4, 5 and 6. There are not many significant differences between these two groups of regressions. The lagged income inequality has a strong positive effect on education spending for both regression groups, confirming our main theoretical prediction. More specifically, a rise of 1% in lagged income inequality has a positive effect of 2.35% on education spending per student when primary and secondary levels are considered together, 3.01% and 2.15% for primary and secondary levels respectively when they are considered separately. Furthermore, both public pensions and old dependency ratio have a positive individual effect on education spending, however their interaction indicates that the effect of ODR becomes negative beyond a certain level of public pensions per pensioner.[21] More specifically, the effect of ODR on primary and secondary

[21] Isolating the interaction effect of the ODR and PubPen on total education spending, we obtain the expression below:

$$\text{EPSPPSE} = 0.0420 \cdot \text{ODR} + 0.0648 \cdot \text{TPS} - 0.0024 \cdot \text{ODR} \cdot \text{PubPen}$$

In order to obtain the effect of the old dependency ratio on total education spending, we take the

Table 4.3.: *Interaction Effect and Education Spending*

	Fixed Effects			Fixed Effects-Driscoll-Kraay Standard Errors		
	(1) ESPSPSE	(2) ESPSPE	(3) ESPSSE	(4) ESPSPSE	(5) ESPSPE	(6) ESPSSE
L.24.Gini	0.0235**	0.0301*	0.0215**	0.0235***	0.0301***	0.0215***
	(0.008)	(0.012)	(0.007)	(0.005)	(0.007)	(0.004)
PubPen	0.0648***	0.0867***	0.0549*	0.0648***	0.0867***	0.0549***
	(0.017)	(0.016)	(0.020)	(0.008)	(0.011)	(0.009)
ODR	0.0420***	0.0487***	0.0380**	0.0420***	0.0487***	0.0380**
	(0.009)	(0.012)	(0.011)	(0.007)	(0.005)	(0.012)
ODR*PubPen	-0.0024**	-0.0033***	-0.0017*	-0.0024***	-0.0033***	-0.0017**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
GDPpc	0.0422***	0.0326***	0.0460***	0.0422***	0.0326***	0.0460***
	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.004)
SHPRPSE	1.2158*			1.2158***		
	(0.448)			(0.289)		
SHPRPE		2.0764†			2.0764**	
		(1.020)			(0.579)	
SHPRSE			1.2190***			1.2190***
			(0.244)			(0.117)
Ctry. & Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	294	315	304	294	315	304
Countries	31	32	33	31	32	33
F-Test	338.74***	97.28***	402.41***	283621.49***	114652.06***	622700.15***
R^2-within	0.8079	0.7924	0.7562	0.8079	0.7924	0.7562

NOTE: Two-way fixed effects regressions with robust standard errors (regression 1 to 3) and Driscoll-Kraay standard errors corrected for heteroscedasticity, autoregressive process of order 2 (regression 4 to 6) reported in parentheses, ***p<0.001, **p<0.01, *p<0.05, †p<0.10. The standard errors are clustered over the number of countries used in each regression. Dependent variables: education spending per student in primary (ESPSPE), secondary (ESPSSE), primary & secondary education (ESPSPSE) are in logs. L.24.Gini: is a lag (24 years) of the Gini index on pre tax and transfers income, ODR: old dependency ratio. Public pensions spending per pensioner (PubPen) and GDPpc are measured in $1,000 PPP (constant 2011). Share of private education in total primary (SHPRPE), secondary (SHPRSE), primary & secondary (SHPRPSE) education. Constant is not reported but included in the above regressions.

education turns to be negative when the level of public pensions per retiree is beyond $14,000 (reg. 5), $22,000 (reg. 6), respectively and $17,000 when considered together (reg. 4). Finally, the share of private education in primary, secondary has a positive impact on education spending just as it is expected by the theory.

The results of Table 4.3 empirically support the theoretical predictions that we examine in this section. Next, we want to investigate the effects of the income inequality and ageing using a dynamic panel approach in order to consider possible path dependence in the determination of education spending.

first derivative of EPSPPSE with respect to the ODR:

$$\partial \text{EPSPPSE}/\partial \text{ODR} = 0.0419 - 0.0023 \cdot \text{PubPen}$$

4.4.3. Dynamic Panel Analysis

So far, it has been implicitly assumed in our model that the past values of the dependent variable do not play any role in the formulation of its current value. However, the current level of education spending might depend on its past levels. Hence, we include as an additional regressor only the first lag of education spending per student. This particular specification of the model implies that we assume that the education spending per student depends on its value in the previous period. Here, we can not apply the previous estimation techniques to the dynamic panel model because the lag of dependent variable is correlated with fixed effects in the error term (dynamic panel bias, see Roodman, 2009).

Moreover, we are not able to exclude the possibility of having endogeneity problems in our previous and current econometric model due to the reverse causality from education spending to fertility and consequently to population ageing (ODR).[22] Also, we can not exclude the possibility of Tiebout effects in the international arena that can influence the fertility rate even at a cross-country level (for a discussion see Persson and Tabellini, 2000). In our case, an example of Tiebout sorting could be the immigration among OECD countries due to better education systems or welfare states. These threats to the internal validity of our model can bring potential biases to our estimations.

In order to address the aforementioned endogeneity concerns and incorporate the lag of the dependent variable as an additional regressor, we employ the "difference GMM" or Arellano-Bond estimation method introduced by Holtz-Eakin et al. (1988) and Arellano and Bond (1991).[23] [24] For this purpose we consider an autoregressive model of 1st order in education spending. We use the following specification:

[22] However, one can argue that this effect is taking place in the long-run. In other words, the age structure if affected is only affected in the long-run and it is fixed and predetermine in the short-run. Also, the impact of education on fertility is far from straightforward. In the past it was thought that more educated women tend to have fewer children (Becker et al., 1990; Galor and Weil, 1996) due to the increasing opportunity cost, however in the most recent study Esping-Andersen and Billari (2015) point to a reversion of this negative relationship.

[23] The Arellano and Bond estimator forms moment conditions using lagged-levels of the dependent variable and the predetermined variables with first-differences of the disturbances. This estimation technique transforms all regressors – by differencing them and removing the fixed effects – and uses Generalised Method of Movements (Hansen, 1982).

[24] When applying Arellano-Bond estimation to the model given by equation (4.31), we classify our regressors with respect to their level of exogeneity. We set as exogenous variables, the lag of income inequality and the private share of enrolments. As predetermined variables we set the public pensions per retiree and ODR. Finally, GDP per capita enters as endogenous variable.

$$\ln(Y_{i,t}) = \gamma \ln(Y_{i,t-1}) + \beta X'_{i,t} + u_i + \delta_t + \varepsilon_{i,t}, \tag{4.31}$$

where $Y_{i,t}$ is public education spending per student of country i at time t, and $Y_{i,t-1}$ is the first lag of public education spending per student. Just as before, the β is a coefficient vector, the u_i is the unobserved country-level effect and δ_t represents the time fixed effects, respectively. Finally, $\varepsilon_{i,t}$ is the idiosyncratic error term. The vector X includes all the regressors used in our estimations.

In Table 4.4 we present the estimations when applying difference GMM to the above specified model. First, in regressions 1, 3 and 5 we estimate the dynamic model without the interaction term between ODR and PubPen. Second, when we include the interaction term – in regressions 2, 4 and 6 – the effect of the lag of education spending is statistically significant and positive. In this case, the coefficients are lower than without the interaction term. More specifically, a one percent increase in education spending of the previous year increases the current spending of total primary ans secondary public education by 0.80% (0.79% and 0.60% in primary and secondary, respectively). However, when we include the interaction term the effect is significantly lower, it is 0.38% for total primary and secondary, 0.32% for only primary and 0.39% for only secondary. One possible explanation for this could be that the interaction effect is absorbed by the lag of education when the interaction of ODR with PubPen is not considered.

Regarding our main explanatory variables, the coefficients have the expected sign, although not all of them are statistically significant. ODR has a negative but non-significant effect on all levels of education spending when we do not take into account its interaction effect with public pensions per pensioner (see regressions 1, 3, and 5). However, when the interaction term is considered the old dependency ratio has a negative impact on primary and secondary education spending per student only when public pensions spending per pensioner is beyond \$14,000 (reg. 2).[25] The same effect is negative when public pensions spending per pensioner is beyond \$17,000 when we consider only primary education, a level considerably higher compared to \$14,000 in regression 4, Table 4.3. The effect of income inequality on education spending is statistically significant and positive (about 1.10-1.18%) for primary education spending per student (regression 3 and 4). However,

[25] The effect of the interaction is determined through the partial derivative just as in the previous section.

Table 4.4.: *Dynamic Panel Estimation*

	(1) ESPSPSE	(2) ESPSPSE	(3) ESPSPE	(4) ESPSPE	(5) ESPSSE	(6) ESPSSE
L.ESPSPSE	0.8013***	0.3808**				
	(0.087)	(0.128)				
L.ESPSPE			0.7990***	0.3252*		
			(0.086)	(0.162)		
L.ESPSSE					0.6082***	0.3959**
					(0.117)	(0.124)
L.24.Gini	0.0037	0.0108	0.0106*	0.0188*	-0.0018	0.0011
	(0.006)	(0.007)	(0.005)	(0.008)	(0.007)	(0.007)
PubPen	-0.0059	0.0748**	-0.0143	0.0846*	0.0182	0.0724*
	(0.010)	(0.028)	(0.011)	(0.038)	(0.013)	(0.033)
ODR	-0.0223	0.0389*	-0.0075	0.0573*	-0.0291	0.0264
	(0.018)	(0.018)	(0.012)	(0.023)	(0.024)	(0.029)
ODR*PubPen		-0.0028**		-0.0035**		-0.0023*
		(0.001)		(0.001)		(0.001)
GDPpc	0.0119	0.0137	0.0108	0.0055	0.0188†	0.0155*
	(0.009)	(0.009)	(0.009)	(0.016)	(0.011)	(0.007)
SHPRPSE	0.4600	0.7748				
	(0.525)	(0.685)				
SHPRPE			0.2931	0.7017		
			(0.494)	(0.881)		
SHPRSE					0.4364	0.8213
					(0.494)	(0.507)
Instruments	73	74	76	77	74	75
Sargan-Test	0.7181	0.6980	0.7626	0.8850	0.1591	0.1351
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	216	216	242	242	225	225
Countries	29	29	31	31	30	30
χ^2 test	1841.77***	2766.40***	19129.95***	2158.69***	1960.11***	9695.37***

Note: One-step GMM estimation, Arellano-Bond robust VCE estimator. Robust standard errors reported in parentheses, ***p<0.001, **p<0.01, *p<0.05, †p<0.10. Time fixed effects included in all regressions. The null hypothesis of the Arellano-Bond test for zero autocorrelation: no autocorrelation, is rejected only at order 1 but not at higher orders. The null hypothesis of the Sargan test of overidentifying restrictions: overidentyfing restrictions are valid, is not rejected. In the specification of the model we use PubPen and ODR as predetermined variables and GDPpc as an endogenous variable. Dependent variables: education spending per student in primary (ESPSPE), secondary (ESPSSE), primary & secondary education (ESPSPSE) are in logs. L.24.Gini: is a lag (24 years) of the Gini index on pre tax and transfers income, ODR: old dependency ratio. Public pensions spending per pensioner (PubPen) and GDPpc are measured in $1,000 PPP (constant 2011). Share of private education in total primary (SHPRPE), secondary (SHPRSE), primary & secondary (SHPRPSE) education. Constant is not reported but included in the above regressions.

the effect is not statistically significant when we consider primary and secondary education jointly (reg.1 and 2). Finally, the effect on secondary education is positive but insignificant.

In our empirical analysis we use two different specifications to estimate the effect of income inequality and population ageing on education spending per student. We can conclude from our baseline specification that there is a positive effect of higher pre-tax and transfers income inequality on education spending per student. When we extend the specification to its dynamic form, we find mixed results regarding the effect of income inequality on education spending per student. More specifically, the effect of income inequality on education spending is mainly driven by the primary education level. Furthermore, the results of both specifications indicate that population ageing has a negative effect on education spending when there is a competition for fiscal resources, namely, pensions spending per pensioner is above a certain level.

4.5. Conclusion

In the recent decades two major trends in income inequality and population ageing have generated significant concerns about the sustainability of the welfare state. The higher income inequality and the increasing elderly population have fuelled the intragenerational and intergenerational conflict, respectively, and in turn have affected the public financing of public education and pensions. The former is a conflict within generation and it is between "rich" and "poor" groups of population over taxation for public provision of pensions and education. The latter conflict is between generations, as young and old have different preferences how to allocate public resources. The aim of this paper is to investigate the effect of these trends on public education and pensions spending per student and retiree, respectively.

To this end we developed a two-dimensional political economy model with public and private education and public pay-as-you-go pension scheme. Our model takes into account both political conflicts and uses the probabilistic voting model to examine the political outcome of the voting process on pensions and education given the preferences of each voting group. Our contribution is to examine those two trends simultaneously in order to understand the mechanisms through which they affect the public finance of education and pensions.

The model predicts that income inequality has a positive impact on education spending per student and the level of pensions per pensioner. This effect goes through the participation in public schooling. An increase in income inequality will increase the share of parents that choose to send their children to private schools, reducing the participation in public schools. Hence, increasing the spending per enrolled student and releasing fiscal resources that can be allocated towards a more generous level of pensions. When the state/government is the main provider of schooling an increase in income inequality would improve both the level of education and pensions and reduce the general tax level. The second theoretical prediction of our model states that a rise in the share of elderly population has a negative effect on education spending per student and worsens the level of pensions that every retiree is entitled to. This outcome is a result of a fiscal leakage that comes along with the rise in the population of elderly and puts more pressure on the welfare state.

Our empirical strategy concentrates on the effect of income inequality on education spending in a majority public education regime. We find support of the theoretical claims using OECD data on pensions and education, inequality and ageing. More specifically, we show evidence of the negative effect of old dependency ratio on education when we take into account that the impact could depend on the level of pensions. However, we obtain mixed results regarding the effect of income inequality on education spending.

An interesting direction for future research could follow an alternative approach by relaxing the assumption of a balanced government budget that we make in this model. The possibility to finance pensions and education by increasing the government's primary deficit could alter the incentives of the voting groups that we consider in this study. Moreover, it would be interesting to develop a model that considers a political process with a dynamic interaction between private savings and a PAYG pension system. Another possible trajectory concerns the weight of political power of different voting groups in policy-making.

5 Conclusion

The increasing income inequality, opening of the labour market for women, and the demographic transition are changing the economies of developed countries. In this thesis I investigate how these demographical and societal changes are affecting the decisions that economic agents make over their life cycle. In particular, I looked at the effects of income inequality, changes in the gender wage gap, and population ageing. The chapters have followed the life of an economic agent, from the decision of the young about educational attainment, over the fertility choices of adults, to the interaction between education and public pensions.

The second chapter of this thesis focused on the effect that income inequality, in particular wage premiums, play for education attainment and intergenerational social mobility. Intergenerational social mobility is negatively related to income inequality. I argue based on the literature on intergenerational social mobility and empirical evidence, that high school education plays an important role for understanding why countries and regions with higher income inequality feature higher levels of income persistence over generations. I contribute to the literature on intergenerational social mobility by developing a model focused on high school education and the opportunity costs of education in terms of work that poor students face.

I argue that these opportunity cost are an important factor in explaining the educational attainment of children from poor families. In an empirical exercise, I showed evidence that this type of opportunity cost are more important than the returns on education in explaining differences in the high school dropout rate between commuting zones in the U. S. I show in the model that a policy decreasing this kind of opportunity cost by increasing the income of young high school graduates paid through higher taxes on high school graduates could improve high school attainment of poor students. This could work through a decrease in taxes of young graduates, or through direct subsidies on the consumption in high school, e. g. through subsidised school meals or extracurricular activity.

In the model, such a policy applied to college education would also increase college graduation rates and decrease the number of college graduates facing binding credit constraints. As student debt is increasing strongly in the U. S. this is a topic of increasing importance. Such a policy could work through governmental loans, tuition waivers and scholarships financed through higher taxes on high income brackets which are majority college graduates.

In the third chapter, I investigate the the decline in fertility in Germany. I show in an empirical decomposition of the completed fertility rate of German women born between 1930 and 1965 that both intensive fertility and childlessness play an important role in the fertility decline. I argue that the quantity/quality trade-off used in the literature on fertility choices cannot explain this due to the essential complementarity between investment into a child's education and having a child. The literature on childlessness so far abstracts from investment into education, which is the main driver of the quantity/quality trade-off.

This chapter aims at reconciling these two strains of the literature. I show in an event study approach that women face a high child penalty in terms of time. The labour income of women declines strongly after the birth of a first child, and this decline is driven by a decline in labour force participation and hours worked. The hourly wage stays constant, while men do not show any negative change in income after childbirth. Based on these findings, I develop a model featuring a quantity/quality trade off and opportunity cost of having children in terms of work time for women. I show in a calibration exercise that this model is able to generate the general trend in fertility and childlessness through a decrease in the gender wage gap. Therefore I argue that the time costs of having children is one of the main drivers of the decline in fertility and the increase in childlessness.

The results suggests that family policies that aim at increasing fertility levels should focus on the time costs of children for women. Policies such as public child care and full day schooling could help decrease the opportunity cost of having children in terms of work time for women, and thus alleviate the effect of increasing female labour market opportunities on fertility.

The forth chapter of my thesis is studying the effect of income inequality and population ageing on public education and pensions. In most OECD countries, there is an increase in income inequality and population ageing. The increase in income inequality is aggravating the conflict between rich and poor over how much to redistribute through the welfare state, whereas the increase in population ageing is intensifying the conflict between young and old about the allocation of public funds between education and pensions.

Most of the literature consider these conflicts in isolation. We argue that it is important to consider the interaction between them and develop a model that accounts for both conflicts. In this overlapping generations model, agents first decide whether or not to send their children into public education and how many children to have. After that, the adults and the old vote on how much to spend on public education and public pensions according to a probabilistic voting model. We show in this model that an increase in income inequality increases per student spending on public education as well as public pensions. An increase in population ageing decreases both the per student spending on public education as well as public pensions. We find empirical evidence for the effect of population ageing on public education using panel data on OECD countries, but we find only mixed evidence on the effect of income inequality.

Concerning the second chapter, an interesting approach for future research would be to introduce heterogeneity in terms of ability. This would allow to investigate the effect of income inequality on the ability distribution of graduates from different parental backgrounds. Furthermore, I think it would be interesting to introduce social mobility into a fertility model as developed in the third chapter and to investigate the interaction between fertility choices and intergenerational social mobility.

With regard to the third chapter, a promising direction for future studies would be to conduct the fertility and child penalty decomposition for different countries and at different stages of economic development. This would allow to test the framework developed in this chapter for its explanatory power in different contexts. Moreover, I think that this framework could be used to analyse the effect of family policies on fertility and childlessness.

With regard to the forth chapter, I think that future research should look at the effect of relaxing the assumption of a balance budget in models of pensions and education and assessing the effect of introducing difference in political power into the voting model. This would allow to evaluate dynamic effects of pensions and generate stronger diverging interests between generations.

The interaction of demographical and economic trends is often not straightforward and understanding the ways in which they interact is important for designing effective policies. The purpose of the studies that I conducted throughout this thesis is to contribute to our understanding of the effect of these societal trends and to contribute to the design of policies oriented to deal with their consequences for the welfare of agents along the life cycle.

A Appendix to Chapter 1

A.1. Detailed Comparative Analysis

A decrease in the opportunity cost of high school education can be caused by an increase in $y_{1,t}^H$ or by a decrease in $y_{1,t}^N$. I first assume that the change in the wage premium is caused by a decrease in $y_{1,t}^N$.

Proposition 13. *A decrease in the income of young high school dropouts $y_{1,t}^N$ decreases b_t^H.*

Proof. The partial derivative of b_t^H with respect to $y_{1,t}^N$ is equal to

$$\frac{\partial b_t^H}{\partial y_{1,t}^N} = -(1-\tau)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\left\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\right.$$
$$\left. - \left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\right\}^{-1},$$

which following Lemma 1 is positive. □

The denominator is equal to the difference $\bar{c}_{1,t}^H - c_{1,t}^{N*}$. Thus the higher the difference $\bar{c}_{1,t}^H - c_{1,t}^{N*}$ and the lower the constraint first period consumption of a high school graduate, the higher will be the decrease in b_t^h. The cut-off value for facing a binding credit constraint $\hat{b}_t^H$ is not affected by a change in $y_{1,t}^N$. Thus if it decreases, more students facing a binding credit constraint will graduate from high school.

I now assume that the change in the opportunity cost of high school education is caused by an increase in $y_{1,t}^H$ and that $y_{1,t}^N$ stays constant.

Proposition 14. *An increase in the income of young high school graduates $y_{1,t}^H$ decreases b_t^H and $\hat{b}_t^H$.*

Proof. The partial derivative of b_t^H with respect to $y_{1,t}^H$ is equal to

$$\frac{\partial b_t^H}{\partial y_{1,t}^H} = (1-\tau)\left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]$$
$$\cdot\left\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\right.$$
$$\left. - \left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\right\}^{-1} \tag{A.1}$$

which following Lemma 1 is negative. One can derive the partial derivative of $\hat{b}_t^H$ with respect to $y_{1,t}^H$

$$\frac{\partial \hat{b}_t^H}{\partial y_{1,t}^H} = \frac{1-\tau}{\Gamma} - \frac{(1-\tau)(1+\Gamma)}{\Gamma}$$
$$= -(1-\tau),$$

which is also negative. □

The denominator of (A.1) is equal to the the difference $\bar{c}_{1,t}^H - c_{1,t}^{N*}$ times $\bar{c}_{1,t}^H$. Thus the larger the difference between constraint and unconstrained consumption and the larger the constraint first period consumption, the smaller the effect of $y_{1,t}^H$ on b_t^H. The nominator is equal to $\bar{c}_{1,t}^H c_{1,t}^{N*}$. The higher $\bar{c}_{1,t}^H$ and $c_{1,t}^{N*}$, the larger is the increase. Thus a increase in $y_{1,t}^H$ increases the value of $\hat{b}_t^H$. It will increase the number of student graduating from high school, and increase the number of students that are not facing a binding credit constraint when graduating from high school.

Thus from a policy perspective, a decrease in $y_{1,t}^N$ might increase the number of high school graduates, but only the number of graduates facing a binding credit constraint, whereas an increase in $y_{1,t}^H$ increases the number of high school graduates and increases the number of high school graduates that are not bound by the credit constraint.

As in the case of the opportunity cost of education, the increase in the return on high school education can come from a decrease in $y_{2,t+1}^N$ or from an increase in $y_{2,t+1}^H$.

Proposition 15. *A decrease in the wages of old high school dropouts $y_{2,t+1}^N$ decreases the transfer b_t^H needed in order to graduate from high school.*

Proof. The partial derivative of b_t^H with respect to $y_{2,t+1}^N$ is equal to

$$\frac{\partial b_t^H}{\partial y_{2,t+1}^N} = -(1-\tau)\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\Big\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]$$
$$-\left[(1-\tau)(Ry_{1,t}^N + y_{2,t}^N) - g_2^N + R(b_t^H - g_1^N)\right]\Big\}^{-1},$$

which following Lemma 1 is positive. □

This expression is increasing in the first period consumption of high school graduates and decreasing in the difference between $\bar{c}^H_{1,t}$ and $c^{N*}_{1,t}$. $y^N_{2,t+1}$ has no effect on $\hat{b}^H_t$, thus a decrease in $y^N_{2,t+1}$ will increase the number of high school graduates that graduate while facing a binding credit constraint.

Proposition 16. *An increase in the income of old high school graduates $y^H_{2,t+1}$ decreases the transfer b^H_t needed in order to graduate from high school but increases the transfer needed in order to not face a binding credit constraint when acquiring high school education $\hat{b}^H_t$.*

Proof. The partial derivative of b^H_t with respect to $y^H_{2,t+1}$ is equal to

$$\begin{aligned}\frac{\partial b^H_t}{\partial y^H_{2,t+1}} = \Gamma &\left[(1-\tau)y^H_{1,t} - g^H_1 + b^H_t\right] \\ &\cdot \left[(1-\tau)(Ry^N_{1,t} + y^N_{2,t}) - g^N_2 + R(b^H_t - g^N_1)\right] \\ &\cdot (y^H_{2,t+1} - g^H_2)^{-1} \left\{(1+\Gamma)R\left[(1-\tau)y^H_{1,t} - g^H_1 + b^H_t\right]\right. \\ &\left. - \left[(1-\tau)(Ry^N_{1,t} + y^N_{2,t}) - g^N_2 + R(b^H_t - g^N_1)\right]\right\}^{-1},\end{aligned}$$

which following Lemma 1 is negative. The partial derivative of $\hat{b}^H_t$ with respect to $y^H_{2,t}$ is equal to

$$\frac{\partial \hat{b}^H_t}{\partial y^H_{2,t+1}} = \frac{1-\tau}{R\Gamma},$$

which is positive. □

This negative effect on b^H_t is the larger, the larger first period income of dropouts and high school graduates, the lower the difference $\bar{c}^H_{1,t} - c^{N*}_{1,t}$ and the lower $y^H_{2,t+1}$ itself is. Thus an increase in $y^H_{2,t+1}$ increases the number of high school graduates, but also increases the number of high school graduates that are facing a binding credit constraint.

The return on college education can increase because the income of old high school graduates decreases, or because the income of college graduates increases. A change in the income of high school graduates has the following direct effect on b^C_t:

Proposition 17. *A decrease in the income of high school graduates $y^H_{2,t+1}$ decreases the transfer b^C_t needed in order to graduate from college.*

Proof. The partial derivative of b_t^C with respect to $y_{2,t+1}^H$ is equal to

$$\frac{\partial b_t^C}{\partial y_{2,t+1}^H} = -\frac{(1-\tau)(b_t^C - g_1^C)}{(1+\Gamma)R(b_t^C - g_1^C) - \left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_1^H + R(b_t^C - g_2^H\right]},$$

which following Lemma 1 is positive. □

A decrease in $y_{2,t+1}^H$ will decrease b_t^C and thus increases the number of college graduates. This increase will be the larger, the larger b_t^C and the smaller $\bar{c}_{1,t}^C - c_{1,t}^{H*}$. As it has no effects on $\hat{b}_t^C$, it will only increase the number of credit constrained college graduates, but not the number of credit-non constrained college graduates.

Proposition 18. *An increase in the income of college graduates $y_{2,t+1}^C$ decreases the transfer b_t^C needed in order to graduate from college and increases the transfer $\hat{b}_t^C$ needed in order to be not facing a binding credit constraint when graduating from college.*

Proof. The partial derivative of b_t^C with respect to $y_{2,t+1}^C$ is equal to

$$\begin{aligned}\frac{\partial b_t^C}{\partial y_{2,t+1}^C} = & \frac{\Gamma(b_t^C - g_1^C)}{y_{2,t+1}^C - g_2^C}\left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_1^H + R(b_t^C - g_2^H\right] \\ & \cdot \left\{(1+\Gamma)R(b_t^C - g_1^C) - \left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H)\right.\right. \\ & \left.\left. -g_1^H + R(b_t^C - g_2^H\right]\right\}^{-1},\end{aligned}$$

which following Lemma 1 is negative. The partial derivative of $\hat{b}_t^C$ with respect to $y_{2,t+1}^C$ is equal to

$$\frac{\partial \hat{b}_t^C}{\partial y_{2,t}^C} = \frac{1-\tau}{R\Gamma},$$

which is positive. □

An increase in $y_{2,t}^C$ decreases b_t^C and thus increases the number of college graduates. The effect is the larger, the larger life time income of high school graduates, the larger b_t^C, and the smaller $\bar{c}_{1,t}^C - c_{1,t}^{H*}$ and $y_{2,t+1}^C$. A change in $y_{2,t+1}^C$ does not only affect b_t^C but also the cut-off value below which college graduates are facing a binding credit constraint. Thus an increase in $y_{2,t+1}^C$ increases the number of college graduates in total, and the number of college graduates facing a binding credit constraint, but decreases the number of college graduates not facing a binding credit constraint.

A.2. Proofs of Policy Analysis

Proof of Proposition 5. The partial derivative of b_t^H with respect to g_1^H is

$$\begin{aligned}\frac{\partial b_t^H}{\partial g_1^H} &= -(1+\Gamma)R\left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\\ &\quad\cdot\left\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\right.\\ &\quad\left.-\left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\right\}^{-1},\end{aligned}$$

which following Lemma 1 is positive. The partial derivative of b_t^H with respect to g_2^H is equal to

$$\begin{aligned}\frac{\partial b_t^H}{\partial g_2^H} &= -\Gamma\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\\ &\quad\cdot\left[(t-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\\ &\quad\left[(1-\tau)y_{2,t+1}^H - g_2^H\right]^{-1}\left\{(1+\Gamma)R\left[(1-\tau)y_{1,t}^H - g_1^H + b_t^H\right]\right.\\ &\quad\left.-\left[(1-\tau)(Ry_{1,t}^N + y_{2,t+1}^N) - g_2^N + R(b_t^H - g_1^N)\right]\right\}^{-1},\end{aligned}$$

which following Lemma 1 is also positive. As

$$\frac{\partial b_t^H}{\partial g_2^H} = \frac{\Gamma}{R(1+\Gamma)}\frac{(1-\tau)y_{1,t}^H - g_1^H + b_t^H}{(1-\tau)y_{2,t+1}^H - g_2^H}\frac{\partial b_t^H}{\partial g_1^H}$$

and according to equation (2.6) it is true that

$$\frac{\Gamma}{R(1+\Gamma)}\frac{(1-\tau)y_{1,t}^H - g_1^H + b_t^H}{(1-\tau)y_{2,t+1}^H - g_2^H} < 1,$$

$\partial b_t^H / \partial g_1^H > \partial b_t^H / \partial g_2^H$ and a simultaneous increase in g_2^H and decrease in g_1^H of equal size decreases b_t^H. As

$$\frac{\partial \hat{b}_t^H}{\partial g_1^H} = \frac{1}{\Gamma} - \frac{1+\Gamma}{\Gamma} = -1,$$

and

$$\frac{\partial \hat{b}_t^H}{\partial g_2^H} = \frac{1}{R\Gamma},$$

this policy also decreases $\hat{b}_t^H$. □

Proof of Proposition 6. The partial derivative of b_t^C with respect to g_1^C is

$$\frac{\partial b_t^C}{\partial g_1^C} = -\frac{(1+\Gamma)R\left[(Ry_{1,t}^H + y_{2,t+1}^H) - g_2^H + R(b_t^C - g_1^H)\right]}{(1+\Gamma)R(b_t^C - g_1^C) - \left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_2^H + R(b_t^C - g_1^H)\right]}$$

which following Lemma 1 is positive. The partial derivative of b_t^C with respect to g_1^C is equal to

$$\begin{aligned}\frac{\partial b_t^C}{\partial g_1^C} = &-\Gamma(b_t^C - g_1^C)\left[(Ry_{1,t}^H + y_{2,t+1}^H) - g_2^H + R(b_t^C - g_1^H)\right] \\ &\cdot \left[(1-\tau)y_{2,t+1}^C - g_2^C\right]^{-1}\left\{(1+\Gamma)R(b_t^C - g_1^C)\right. \\ &\left.- \left[(1-\tau)(Ry_{1,t}^H + y_{2,t+1}^H) - g_2^H + R(b_t^C - g_1^H)\right]\right\}^{-1},\end{aligned}$$

which following Lemma 1 is also positive. It follows from this that

$$\frac{\partial b_t^C}{\partial g_2^C} = \frac{\Gamma}{R(1+\Gamma)}\frac{b_t^C - g_1^C}{(1-\tau)y_{2,t+1}^H - g_2^C}\frac{\partial b_t^C}{\partial g_1^C}.$$

According to equation (2.7) it is true that

$$\frac{\Gamma}{R(1+\Gamma)}\frac{b_t^C - g_1^C}{(1-\tau)y_{2,t+1}^H - g_2^C} < 1.$$

Thus $\partial b_t^C/\partial g_1^C > \partial b_t^C/\partial g_2^C$ and a simultaneous increase in g_2^C and decrease in g_1^C of the same size decreases b_t^C. As

$$\frac{\partial \hat{b}_t^C}{\partial g_1^C} = \frac{1}{\Gamma} - \frac{1+\Gamma}{\Gamma} = -1,$$

and

$$\frac{\partial \hat{b}_t^C}{\partial g_2^C} = \frac{1}{R\Gamma},$$

this policy also decreases $\hat{b}_t^C$. □

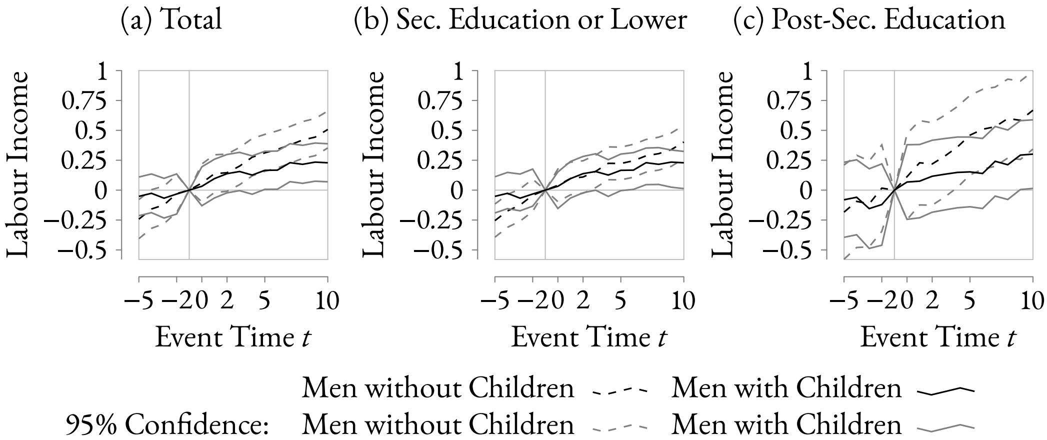
B Appendix to Chapter 2

B.1. Difference in Difference Estimation

I follow Kleven et al. (2018) in estimating the effect of childbirth on net labour income, labour force participation, hours worked, and hourly wage for men and women relative to those that stay childless. For this, I restrict my sample to those for which I have observations above the age of 40, as this is generally considered to be the oldest age for a first birth. This assumption is less reasonable for men than for women, but as seen in the estimation based on only mothers and fathers, the child penalty is only present for women and I am therefore mainly interested in the effect of childbirth on women.

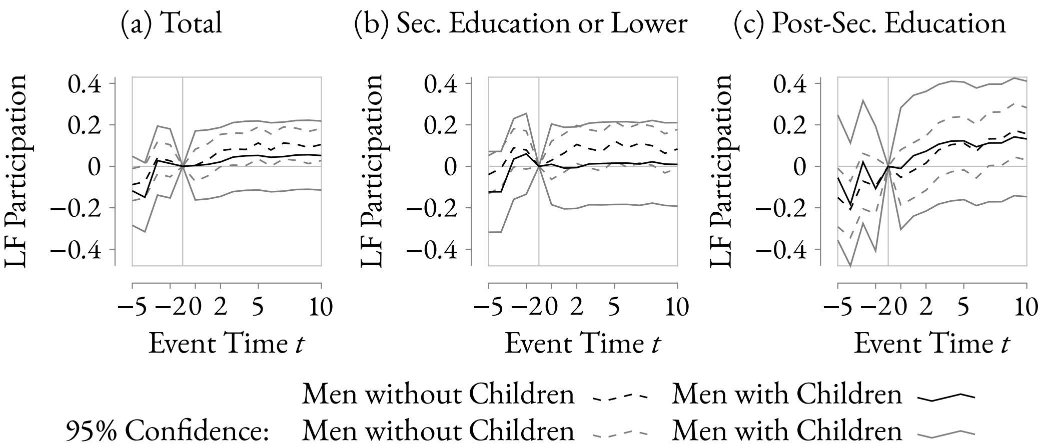
I estimate a log normal distribution on age at first birth for the men and women that gave birth, and then assign placebo first births to the women and men that are childless. As in the event study approach, I restrict the sample to those for which I can observe the 5 years prior and the 10 years after the birth or placebo birth. By doing so I get a sample of 595 women and 666 men. I estimate now the child penalty using a difference in difference estimation. For this difference in difference estimation, age and year fixed effects are not needed. The figures below show the child penalty comparing men that have at least one child to men that have no children, and women that have at least one child to women that are childless. As before, the coefficients are scaled such that they represent the values relative to the value at the year prior to the first birth.

Figure B.1.: *Child Penalty Men Net Labour Income – Diff-In-Diff*



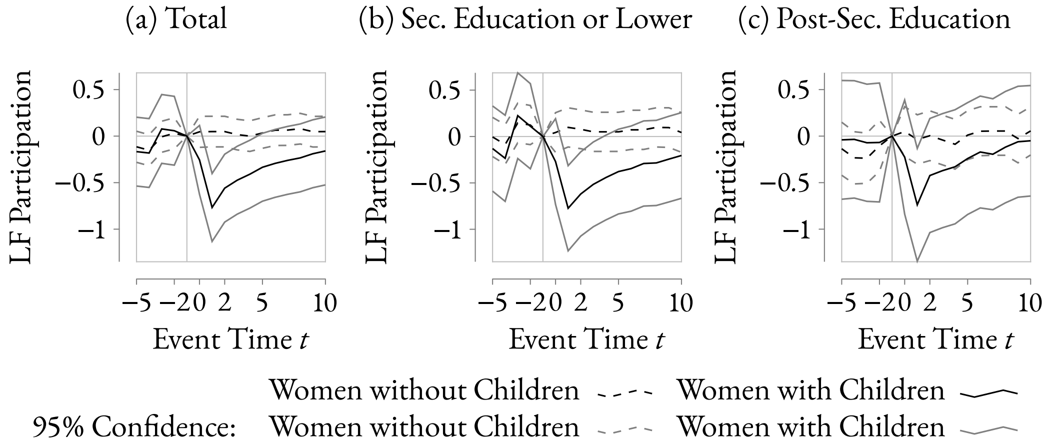
Note: Net labour income of men relative to $t-1$, the year before the birth of the first child. For childless men, placebo births are assigned according to distribution of age at first birth of men with children. Data source: SOEP.

Figure B.2.: *Child Penalty Men Participation – Diff-In-Diff*



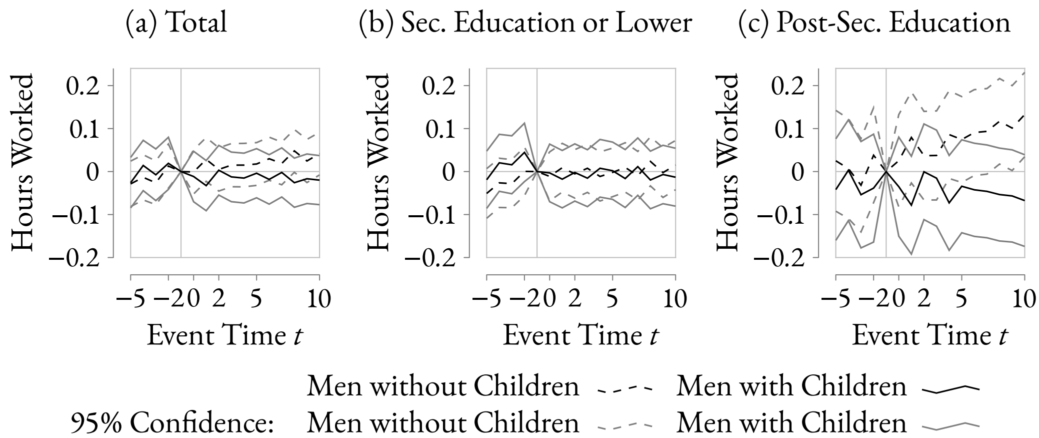
Note: Labour force participation of men relative to $t-1$, the year before the birth of the first child. For childless men, placebo births are assigned according to distribution of age at first birth of men with children. Data source: SOEP.

Figure B.3.: *Child Penalty Women Participation – Diff-In-Diff*



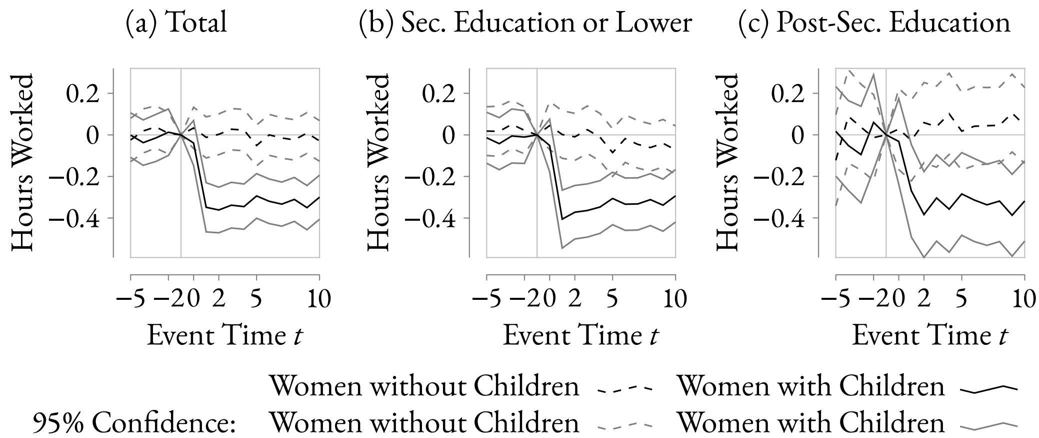
Note: Labour force participation of women relative to $t-1$, the year before the birth of the first child. For childless women, placebo births are assigned according to distribution of age at first birth of women with children. Data source: SOEP.

Figure B.4.: *Child Penalty Men Hours Worked – Diff-In-Diff*



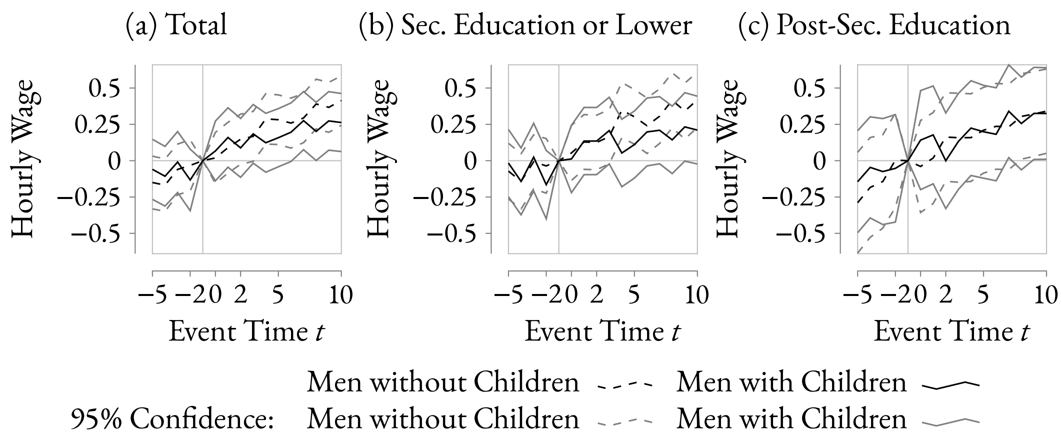
Note: Hours worked conditional on labour force participation relative to $t-1$, the year before the birth of the first child. For childless men, placebo births are assigned according to distribution of age at first birth of men with children. Data source: SOEP.

Figure B.5.: *Child Penalty Women Hours Worked – Diff-In-Diff*



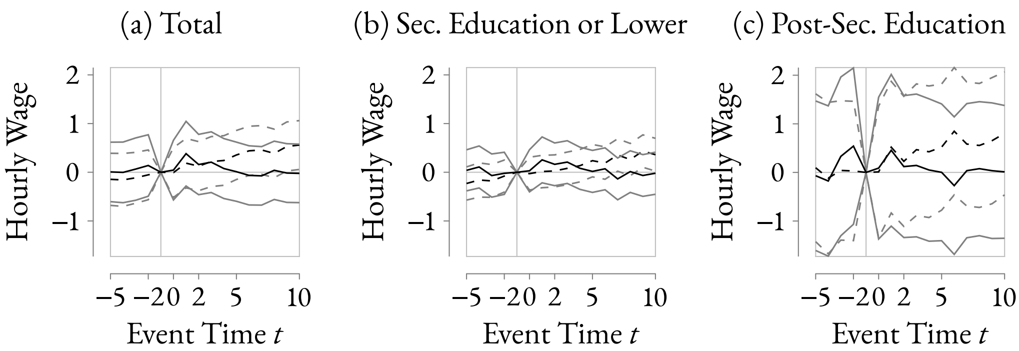
Note: Hours worked conditional on labour force participation relative to $t-1$, the year before the birth of the first child. For childless women, placebo births are assigned according to distribution of age at first birth of women with children. Data source: SOEP.

Figure B.6.: *Child Penalty Men Net Hourly Wage – Diff-In-Diff*



Note: Hourly wage conditional on labour force participation relative to $t-1$, the year before the birth of the first child. For childless men, placebo births are assigned according to distribution of age at first birth of men with children. Data source: SOEP.

Figure B.7.: *Child Penalty Women Net Hourly Wage – Diff-In-Diff*



Women without Children - - - - Women with Children ——
95% Confidence: Women without Children - - - - Women with Children ——

Note: Hourly wage conditional on labour force participation relative to $t - 1$, the year before the birth of the first child. For childless women, placebo births are assigned according to distribution of age at first birth of women with children. Data source: SOEP.

C Appendix to Chapter 3

C.1. The Voting Mechanism

We extend the probabilistic voting model used in de la Croix and Doepke (2009) by introducing the dimension of pensions in the voting process. Hence, voters decide about the tax rate v_t the per student spending on public education s_t, and the per pensioner pension p_t according to a probabilistic voting mechanism based on Lindbeck and Weibull (1987) and Persson and Tabellini (2000). This voting works in the following way: There are two political platforms a and b competing for the votes of the agents. They are competing by offering a policy consisting of a tax rate v_t, a per pensioner pension p_t and a per student education spending s_t that are fulfilling the government budget constraint

$$\int_0^{\tilde{x}_t} s_t n^s g(x)dx + \frac{1}{1+\varrho_{t-1}} p_t = v_t \left\{ \int_0^{\tilde{x}_t} x(1-\phi n^s) g(x)dx \right. \\ \left. + \int_{\tilde{x}_t}^{\infty} \left[x(1-\phi n^e) - e_t^e(x) n^e \right] g(x)dx \right\}.$$

Voters are more likely to vote for the platform that yield them a higher utility. In contrast to the median voter theory, voters do not vote with probability one for the platform that maximises their utility but the probability of voting for platform a instead of platform b is an increasing and differentiable cumulative distribution function on the utility difference between policy a and policy b:

$$F\left\{ U_t\left[x, s_t^a, p_t^a, v_t(s_t^a, p_t^a)\right] - U_t\left[x, s_t^b, p_t^b, v_t(s_t^b, p_t^b)\right] \right\}.$$

This means that the voting decision is not discrete but rather a continuous function of the policy offered by both parties. The uncertainty of the voting is the result of the presence of ideological bias which is independent of the proposed policies. From this follows that the political platforms do not only appeal to the median voter, but consider the preferences of all voters instead. This allows us to aggregate the preferences of different demographical groups (rich, poor, young and old) in the policy function, which leads to the following objective function:

$$\Omega(s_t, p_t) = \int_0^{\tilde{x}_t} U_t^s(x, s_t, p_t, v_t(s_t, p_t))g(x)dx + \int_{\tilde{x}_t}^{\infty} U_t^e(x, s_t, p_t, v_t(s_t, p_t))g(x)dx + \frac{1}{1+\varrho_{t-1}} U_t^o(p_t).$$

Both parties maximise their expected vote share in a symmetrical way, leading to an equilibrium where both political platforms converge to the same policy $\{v_t^*, s_t^*, p_t^*\}$. The equilibrium policy is the policy that maximises the objective function above.

C.2. Education Regimes

In a majority private education regime with $\Psi_t < 1/2$, participation in public education Ψ_t and the tax rate v_t^* are increasing with income inequality σ and the quality of public education s_t^* and the pensions per pensioner p_t^* are decreasing with σ. In an equally separated education regime, participation in public education, tax rate, quality of public education, and pensions are not affected by changes in inequality. This follows from the proof of Proposition 11, where the first derivative of Ψ_t with respect to σ

$$\frac{\partial \Psi_t}{\partial \sigma} = \frac{\sigma - \left[\frac{1-\eta}{\hat{\eta}\phi\eta}\mathbb{E}_t(s_t) - (1-\sigma)\right]}{2\sigma^2} = \frac{1}{\sigma}\left(\frac{1}{2} - \Psi_t\right)$$

is positive for $\Psi_t < 1/2$, and equal to 0 for $\Psi_t = 1/2$. Following Lemma 2 this means that p_t^* and s_t^* are decreasing in σ and v_t^* is increasing in σ for $\Psi_t < 1/2$ and they are not affected by a change in σ is $\Psi_t = 1/2$.

The mechanism of the effect of an increase in income inequality is the following: an increase in income inequality is decreasing the income of the marginal agent that is indifferent between private and public education if this agent has a below average income. This means that this agent now prefers public education. This increase in public education increases the share of voters with children in public education, but it also increases the number of children in public education. Therefore the total spending on public education increases, but the spending per child decreases. Overall this leads to a decrease in public education quality. The increase in total education spending leads to a decrease in pensions and to an increase in taxes.

C.3. Analysis on the Total Education Spending

In Table C.1 we consider the effect of income inequality and population ageing on total education spending as percentage of GDP. In this specification of the empirical model we use as control variables the level of public pensions, GDP per capita, the share of students in private education and number of students in public primary, secondary and total education. As we can observe, income inequality has a positive effect on primary and secondary education in both specifications of the model. Regarding the non-dynamic panel model in regressions 1,2 and 3, we observe that a percentage rise in past income inequality increases primary total education spending by 0.0325%, secondary by 0.0295%, and the aggregate spending on primary and secondary education by 0.0675%. Old dependency ratio and public pensions per pensioner have positive effect on most levels of education spending considered in the Table C.1.

In regressions 4, 5 and 6 with dynamic panel specification, one percentage increase in income inequality in the past has an impact of about 0.0588% on total education spending (primary and secondary considered jointly), 0.0276% on primary and 0.0284% on secondary total spending. Moreover, our proxy for population ageing (ODR) has a negative but insignificant impact on education spending.

Table C.1.: Total Spending in Primary and Secondary Education as % of GDP

	Fixed effects SCC			*Arellano-Bond*		
	(1) GEPSE	(2) GEPE	(3) GESE	(4) GEPSE	(5) GEPE	(6) GESE
L.GEPSE				0.6107**		
				(0.191)		
L.GEPE					0.6813***	
					(0.125)	
L.GESE						0.3186*
						(0.139)
L.24.Gini	0.0675**	0.0325***	0.0295*	0.0588***	0.0276*	0.0284*
	(0.021)	(0.005)	(0.012)	(0.016)	(0.011)	(0.014)
ODR	0.0392*	-0.0151	0.0312**	-0.0018	-0.0464	0.0246
	(0.015)	(0.009)	(0.008)	(0.067)	(0.032)	(0.039)
PubPen	0.0681***	0.0294**	0.0351**	-0.0191	-0.0503*	0.0375
	(0.016)	(0.008)	(0.010)	(0.042)	(0.020)	(0.024)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Instruments				77	78	77
Sargan-Test				0.0978	0.0972	0.0564
F-test	396921.96***	71246.01***	522487.65***			
Country FE	Yes	Yes	Yes			
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	294	315	304	230	252	238
Countries	31	32	33	29	31	30
R^2	0.3493	0.3732	0.3213			
χ^2				403.79***	659.11***	453.21***

Note: Regressions 1,2 and 3: Fixed effects with robust Driscoll-Kraay standard errors corrected for heteroscedasticity, autoregressive process of order 2. Regressions 4, 5 and 6: One-step GMM estimation, Arellano-Bond robust VCE estimator. Robust standard errors for both groups of regressions are reported in parentheses, ***p<0.001, **p<0.01, *p<0.05, †p<0.10. Time fixed effects included in all regressions. The null hypothesis of the Arellano-Bond test for zero autocorrelation: no autocorrelation, is rejected only at order 1 but not at higher orders. The null hypothesis of the Sargan test of overidentifying restrictions: overidentyfing restrictions are valid, is not rejected. In the specification of the model we use PubPen and ODR as predetermined variables and GDPpc as an endogenous variable. Dependent variable: total education spending in primary (GEPE), secondary (GESE), primary & secondary education (GEPSE). L.24.Gini: is a lag (24 years) of the Gini index on pre tax and transfers income, ODR: old dependency ratio. Public pensions spending per pensioner (PubPen) and GDPpc are measured in $1,000 PPP (constant 2011). As control variables (not reported) we use the GDPpc, the share of private education in total primary (SHPRPE), secondary (SHPRSE), primary & secondary(SHPRPSE) education, and the number of student in public primary (ENPUBPE), secondary (ENPUBSE) and total primary and secondary (ENPUBPSE) education. Constant is not reported but included in the above regressions.

C.4. Appendix Tables

Table C.2.: *Partial Correlations: Education, Pensions, and ODR*

Variables	ESPSPSE	PubPen	ODR
ESPSPSE	1.0000		
Obs	420		
PubPen	0.7334* (0.0000)	1.0000	
Obs	389	803	
ODR	0.4525* (0.000)	0.4606* (0.0000)	1.0000
Obs	420	803	1225

Table C.3.: *Data: Definitions and Sources*

Variable	Definition & Source
ESPSPE, ESPSSE, ESPSPSE	Education spending per enrolled student in primary, secondary, total primary and secondary educational level. It is calculated using the total public education spending and enrollments, Expenditure on Education, UNESCO.
ENPUBPE, ENPUBSE, ENPUBPSE	Enrollments (number of students) in primary, secondary, total primary and secondary educational level (as a % of total (private and public) primary & secondary), Enrollment by type of institution, UNESCO.
SHPRPE, SHPRSE, SHPRPSE	Share of enrollments in private primary & secondary education, World Bank Data: World Development Indicators.
GINI	Gini index of market income inequality before taxes and transfers, The Standardized World Income Inequality Database.
ODR (ODR(20-54))	Old dependency ratio, population over 65(55) years old as % of working age population 20-64(54) years old, World Population Prospects, United Nations
PubPen	Public pensions spending per retiree, calculated using Total Public Pensions as % of GDP and population over 65 years old , Social Expenditure, OECD.
GDPpc	GDP per capita based on purchasing power parity (PPP), World Bank Data: World Development Indicators.

Table C.4.: *Alternative Old Dependency ratio 20-54*

	Fixed Effects (SCC)			*Arellano-Bond*		
	(1) ESPSPSE	(2) ESPSPE	(3) ESPSSE	(4) ESPSPSE	(5) ESPSPE	(6) ESPSSE
L.ESPSPSE				0.4382***		
				(0.126)		
L.ESPSPE					0.4101**	
					(0.158)	
L.ESPSSE						0.4169***
						(0.097)
L.24.GINI	0.0214***	0.0270***	0.0194***	0.0089	0.0163**	0.0009
	(0.005)	(0.006)	(0.004)	(0.006)	(0.006)	(0.006)
ODR(20-54)	0.0255***	0.0283***	0.0235*	0.0211*	0.0303*	0.0117
	(0.005)	(0.005)	(0.009)	(0.010)	(0.012)	(0.014)
PubPen	0.0585***	0.0741***	0.0539***	0.0538**	0.0550*	0.0541*
	(0.008)	(0.010)	(0.009)	(0.021)	(0.024)	(0.025)
ODR(20-54)*PubPen	-0.0016***	-0.0020***	-0.0013**	-0.0014**	-0.0017**	-0.0012+
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Instruments				74	77	75
Sargan-Test				0.5828	0.8022	0.0812
F-test	322930.89***	149133.05***	635177.87***			
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	294	315	304	216	242	225
Countries	31	32	33	29	31	30
R^2	0.8173	0.8011	0.7620			
χ^2				1678.31***	3297.80***	4231.70***

NOTE: Regressions 1,2 and 3: Fixed effects with robust Driscoll-Kraay standard errors corrected for heteroscedasticity, autoregressive process of order 2. Regressions 4, 5 and 6: One-step GMM estimation, Arellano-Bond robust VCE estimator. Robust standard errors for both groups of regressions are reported in parentheses, ***p<0.001, **p<0.01, *p<0.05, †p<0.10. Time fixed effects included in all regressions. The null hypothesis of the Arellano-Bond test for zero autocorrelation: no autocorrelation, is rejected only at order 1 but not at higher orders. The null hypothesis of the Sargan test of overidentifying restrictions: overidentyfing restrictions are valid, is not rejected. In the specification of the model we use PubPen and ODR as predetermined variables and GDPpc as an endogenous variable. Dependent variable: education spending per student in primary (ESPSPE), secondary (ESPSSE), primary & secondary education (ESPSPSE) is in logs. L.24.Gini: is a lag (24 years) of the Gini index on pre tax and transfers income, ODR(20-54): old dependency ratio, people over 55 years old as a percentage of people 20 to 54 years old. Public pensions spending per pensioner (PubPen) and GDPpc are measured in $1,000 PPP (constant 2011). As controls variables (not reported) we use the GDPpc, the share of private education in total primary (SHPRPE), secondary (SHPRSE), primary & secondary (SHPRPSE) education. Constant is not reported but included in the above regressions.

Bibliography

Aaronson, Daniel and Bhashkar Mazumder, "Intergenerational Economic Mobility in the United States, 1940 to 2000," *The Journal of Human Resources*, 2008, *43* (1), 139–172.

__ , Fabian Lange, and Bhashkar Mazumder, "Fertility Transitions Along the Extensive and Intensive Margins," *The American Economic Review*, 2014, *104* (11), 3701–3724.

Abío, Gemma, Géraldine Mahieu, and Concepció Patxot, "On the Optimality of PAYG Pension Systems in an Endogenous Fertility Setting," *Journal of Pension Economics and Finance*, 2004, *3* (1), 35–62.

Acemoglu, Daron and David Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Vol. 4b, Elsevier Inc., 2011, pp. 1043–1171.

Adda, Jérôme, Christian Dustmann, and Katrien Stevens, "The Career Costs of Children," *Journal of Political Economy*, 2017, *125* (2), 293–337.

Alders, Peter and D. Peter Broer, "Ageing, fertility, and growth," *Journal of Public Economics*, 2005, *89* (5-6), 1075–1095.

Alonso-Carrera, Jaime, Jordi Caballé, and Xavier Raurich, "Fiscal policy, composition of intergenerational transfers, and income distribution," *Journal of Economic Behavior and Organization*, 2012, *84* (1), 62–84.

__ , Xavier Raurich, and Jordi Caballé, "Intergenerational Mobility, Occupational Decisions and the Distribution of Wages," 2016.

Antonczyk, Dirk, Thomas DeLeire, and Bernd Fitzenberger, "Polarization and Rising Wage Inequality: Comparing the US and Germany," 2010.

Arellano, Manuel and Stephen Bond, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies*, 1991, *58* (2), 277.

Autor, David H and David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *The American Economic Review*, 2013, *103* (5), 1553–1597.

Bailey, Martha J and Susan M Dynarski, "Inequality in Postsecondary Education," in Greg Duncan and Richard J. Murnane, eds., *Whither Opportunity? Rising Inequality, Schools and Children's Life Chances*, 2011, chapter 6, pp. 118–131.

Baudin, Thomas, David de la Croix, and Paula E. Gobbi, "Fertility and childlessness in the United States," *American Economic Review*, 2015, *105* (6), 1852–1882.

_ , _ , and _ , "Endogenous Childlessness and Stages of Development," *Journal of the European Economic Association*, 2018, *00* (0), 1–59.

_ , _ , and _ , "Childlessness and Economic Development: a Survey," in A. Bucci, A. Prskawetz, and K. Prettner, eds., *Human Capital and Economic Growth: The Impact of Health, Education and Demographic Change*, Palgrave - MacMillan, 2019.

Baum, Christopher F., "Residual diagnostics for cross-section time series regression models," *The Stata Journal*, 2001, *1* (1), 101–104.

Bearse, Peter, Gerhard Glomm, and Debra Moore Patterson, "Endogenous Public Expenditures on Education," *Journal of Public Economic Theory*, 2005, *7* (4), 561–577.

Becker, Gary S., "An Economic Analysis of Fertility," in Ansley J. Coale, ed., *Demographic and Economic Change in Developed Countries*, Princton, NJ: Princeton University Press, 1960, pp. 209 – 240.

_ and Nigel Tomes, "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 1986, *4* (3), 1–39.

_ and Robert J. Barro, "A Reformulation of the Economic Theory of Fertility.," *Quarterly Journal of Economics*, 1988, *103* (1), 1–25.

_ , Kevin M. Murphy, and Robert Tamura, "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy*, 1990, *98* (5), 12–37.

Bisin, Alberto and Thierry Verdier, "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, 2001, *97* (2), 298–319.

Blackburn, Keith and Giam Pietro Cipriani, "Intergenerational transfers and demographic transition," *Journal of Development Economics*, 2005, *78* (1), 191–214.

Boadway, Robin W. and David E. Wildasin, "A Median Voter Model of Social Security," *International Economic Review*, 1989, *30* (2), 307–328.

Boldrin, Michele, Larry Jones, and Alice Schoonbroodt, "From Busts to Booms, in Babies and Goodies," 2005.

Booth, Alison L. and Hiau Joo Kee, "Intergenerational Transmission of Fertility Patterns in Britain," 2006.

Breyer, Friedrich and Ben Craig, "Voting on social security: Evidence from OECD countries," *European Journal of Political Economy*, 1997, *13* (4), 705–724.

_ and Klaus Stolte, "Demographic change, endogenous labor supply and the political feasibility of pension reform," *Journal of Population Economics*, 2001, *14* (3), 409–424.

Browning, Edgar K., "Why the Social Insurance Budget Is Too Large in a Democracy," *Economic Inquiry*, 1975, *13* (3), 373–388.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics Using Stata: Revised Edition*, A Stata College Station, Texas : Stata Press, 2010.

Casamatta, G. and L. Batté, "The Political Economy of Population Aging," in John Piggott and Alan Woodland, eds., *Handbook of the Economics of Population Aging, vol 1*, Elsevier, 2016, chapter 7, pp. 381–444.

Castles, Francis G., "On religion and public policy: Does Catholicism make a difference?," *European Journal of Political Research*, 1994, *25* (1), 19–40.

Cattaneo, Alejandra M. and Stefan C. Wolter, "Are the elderly a threat to educational expenditures?," *European Journal of Political Economy*, 2009, *25* (2), 225–236.

Checchi, Daniele, Andrea Ichino, and Aldo Rustichini, "More equal but less mobile? Education financing and intergenerational mobility in Italy and in the US," *Journal of Public Economics*, 1999, *74*, 351–393.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

_ , _ , _ , _ , and Nicholas Turner, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economic Review Papers and Proceedings*, 2014, *104* (5), 141–147.

Corak, Miles, "Do Poor Children Become Poor Adults? Lessons from a Cross-Country Comparison of Generational Earnings Mobility," *Research on Economic Inequality*, 2006, *13* (1993), 143–188.

_ , "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 2013, *27* (3), 79–102.

_ , "Inequality from Generation to Generation: The United States in Comparison," in Robert S. Rycroft, ed., *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*, 2013, chapter 6.

_ , Matthew J. Lindquist, and Bhashkar Mazumder, "A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States," *Labour Economics*, 2014, *30*, 185–200.

Couch, Kenneth A. and Dean R. Lillard, "Non-linear patterns of intergenerational mobility in Germany and the United States," in Miles Corak, ed., *Generational Income Mobility in North America and Europe*, 2004, chapter 8, pp. 190–206.

de la Croix, David and Matthias Doepke, "Inequality and Growth: Why Differential Fertility Matters," *The American Economic Review*, 2003, *93* (4), 1091–1113.

_ and _ , "To segregate or to integrate: Education politics and democracy," *Review of Economic Studies*, 2009, *76* (2), 597–628.

Deschênes, Olivier, "Unobserved Ability, Comparative Advantage, and the Rising Return to Education in the United States 1970-2000," 2001.

Disney, Richard, "Population ageing and the size of the welfare state: Is there a puzzle to explain?," *European Journal of Political Economy*, 2007, *23* (2), 542–553.

Downs, Anthony, "An Economic Theory of Political Action in a Democracy," *Journal of Political Economy*, 1957, *65* (2), 135–150.

Driscoll, John C. and Aart C. Kraay, "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics*, 1998, *80* (4), 549–560.

Eckstein, Zivi and Kenneth I. Wolpin, "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 1999, *67* (6), 1295–1339.

Esping-Andersen, Gøsta and Francesco C. Billari, "Re-theorizing Family Demographics," *Population and Development Review*, 2015, *41* (1), 1–31.

Fernández, Raquel and Richard Rogerson, "Sorting and Long-Run Inequality," *The Quarterly Journal of Economics*, 2001, *116* (4), 1305–1341.

Galor, Oded and Daniel Tsiddon, "Technological Progress and Economic Growth," *The American Economic Review*, 1997, *87* (3), 363–382.

_ and David N. Weil, "The Gender Gap, Fertility, and Growth," *American Economic Review*, 1996, *86* (3), 374–387.

_ and _ , "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *The American Economic Review*, 2000, *90* (4), 806–828.

_ and Hyoungsoo Zang, "Fertility, income distribution and economic growth: Theory and cross-country evidence," *Japan and the world economy*, 1997, *9* (2), 197–229.

_ and Joseph Zeira, "Income Distribution and Macroeconomics," *The Review of Economics Studies*, 1993, *60* (1), 35–52.

_ and Omer Moav, "Ability-biased technological transition, wage inequality, and economic growth," *Quarterly Journal of Economics*, 2000, *115* (2), 469–497.

_ and _ , "Das Human Kapital: A Theory of the Demise of the Class Structure," *Review of Economic Studies*, 2006, *73*, 85–117.

Glomm, Gerhard and B. Ravikumar, "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality," *The Journal of Political Economy*, 1992, *100* (4), 818–834.

Gobbi, Paula E., "A model of voluntary childlessness," *Journal of Population Economics*, 2013, *26* (3), 963–982.

Goos, Maarten, Alan Manning, and Anna Salomons, "Job polarization in Europe," *The American Economic Review: Papers and Proceedings*, 2009, *99* (2), 58–63.

Gradstein, Mark and Michael Kaganovich, "Aging population and education finance," *Journal of Public Economics*, 2004, *88* (12), 2469–2485.

Greenwood, Jeremy, Nezih Guner, and John A. Knowles, "More on Marriage, Fertility, and the Distribution of Income," *International Economic Review*, 2003, *44* (3), 827–862.

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50* (4), 1029.

Hassler, John and José V. Rodríguez Mora, "Intelligence, Social Mobility and Growth," *The American Economic Review*, 2000, *90* (4), 888–908.

_ , _ , and Joseph Zeira, "Inequality and mobility," *Journal of Economic Growth*, 2007, *12* (3), 235–259.

Hausman, Jerry A., "Specification Tests in Econometrics," *Econometrica*, 1978, *46* (6), 1251–1271.

Hertz, Tom, "Trends in the Intergenerational Elasticity of Family Income in the United States," *Industrial Relations A Journal of Economy and Society*, 2007, *46* (1), 22–50.

Hoechle, Daniel, "Robust standard errors for panel regressions with cross-sectional dependence," *The Stata Journal*, 2007, *7* (3), 281–312.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 1988, *56* (6), 1371.

Ichino, Andrea, Loukas Karabarbounis, and Enrico Moretti, "The Political Economy of Intergenerational Income Mobility," *Economic Inquiry*, 2011, *49*(1), 47–69.

Jäntti, Markus, Bernt Bratsberg, Knut Røed, Oddbjørn Raaum, Eva Österbacka, Anders Björklund, and Tor Eriksson, "American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States," *IZA Discussion Paper*, 2006, (1938), 1–40.

Kaganovich, Michael and Itzhak Zilcha, "Pay-as-you-go or funded social security? A general equilibrium comparison," *Journal of Economic Dynamics and Control*, 2012, *36* (4), 455–467.

Kemnitz, Alexander, "Social security, public education, and growth in a representative democracy," *Journal of Population Economics*, 2000, *13* (3), 443–462.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence From Denmark," 2018.

Konrad, Kai A., "Social security and strategic inter-vivos transfers of social capital," *Journal of Population Economics*, 1995, *8* (3), 315–326.

Kornrich, Sabino and Frank Furstenberg, "Investing in Children: Changes in Parental Spending on Children, 1972-2007," *Demography*, 2013, *50* (1), 1–23.

Kremer, Michael and Daniel L. Chen, "Income distribution dynamics with endogenous fertility," *Journal of Economic Growth*, 2002, *7*(3), 227–258.

Krieger, Tim and Jens Ruhose, "Honey, I shrunk the kids' benefits – revisiting intergenerational conflict in OECD countries," *Public Choice*, 2013, *157* (1-2), 115–143.

Lancia, Francesco and Alessia Russo, "Public Education and Pensions in Democracy: a Political Economy Theory," *Journal of the European Economic Association*, 2016, *14* (5), 1038–1073.

Lee, Chul-In and Gary Solon, "Trends in Intergenerational Income Mobility," 2008.

Lemieux, Thomas, "Postsecondary Education and Increasing Wage Inequality," *The American Economic Review*, 2006, *96* (2), 195–199.

Levy, Gilat, "The Politics of Public Provision of Education," *The Quarterly Journal of Economics*, 2005, *120* (4), 1507–1534.

Lindbeck, Assar and Jörgen W. Weibull, "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, 1987, *52* (3), 273–297.

Lindley, Joanne and Stephen Machin, "Rising Wage Inequality and Postgraduate Education," 2011.

_ and _ , "The Quest for More and More Education: Implications for Social Mobility," *Fiscal Studies*, 2012, *33* (2), 265–286.

Manning, Alan, Marten Maarten Goos, and Alan Manning, "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, 2007, *89* (1), 118–133.

Mayer, Susan E., "How Did the Increase in Economic Inequality between 1970 and 1990 Affect Children's Educational Attainment?," *American Journal of Sociology*, 2002, *107* (1), 1–32.

Mazumder, Bhashkar, "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data," *The Review of Economics and Statistics*, 2005, *87* (2), 235–255.

McNeal, R. B., "Are students being pulled out of High School? The effect of adolescent employment on dropping out," *Sociology of Education*, 1997, *70* (3328), 206–220.

Michailidis, Gianko and Concepció Patxot, "Political viability of public pensions and education. An empirical application," *Applied Economics Letters*, 2018, *26* (03), 245–249.

_ , _ , and Meritxell Solé, "Do pensions foster education? An empirical perspective," *Applied Economics*, mar 2019, pp. 1–24.

Moav, Omer and Oded Galor, "From Physical to Human Capital Accumulation," *The Review of Economic Studies*, 2004, *71* (4), 1001–1026.

Naito, Katsuyuki, "Two-sided intergenerational transfer policy and economic development: A politico-economic approach," *Journal of Economic Dynamics and Control*, 2012, *36* (9), 1340–1348.

OECD, "Pensions at a Glance 2017: OECD and G20 Indicators," Technical Report 2017.

Ono, Tetsuo and Yuki Uchida, "Pensions, education, and growth: A positive analysis," *Journal of Macroeconomics*, 2016, *48*, 127–143.

Pereira, Pedro Telhado and Santiago Budría, "Educational Qualifications and Wage Inequality: Evidence for Europe," 2005.

Persson, Torsten and Guido Enrico Tabellini, *Political economics : explaining economic policy*, MIT Press, 2000.

Pesaran, M. H., "General Diagnostic Tests for Cross Section Dependence in Panels," *Cambridge Working Papers in Economics*, 2004.

Piketty, Thomas, "Social Mobility and Redistributive Politics," *The Quarterly Journal of Economics*, 1995, *110* (3), 551–584.

_ , *Le capital au XXIe siècle*, Seuil, 2013.

Pogue, Thomas F. and L. G. Sgontz, "Social security and investment in human capital," *National Tax Journal*, 1977, *30*, 157–169.

Poterba, James M., "Demographic Structure and the Political Economy of Public Education," *Journal of Policy Analysis and Management*, 1997, *16* (1), 48–66.

Putnam, Robert D., Carl B. Frederick, and Kaisa Snellman, "Growing Class Gaps in Social Connectedness among American Youth," 2012.

Ramey, Garey and Valerie A. Ramey, "The Rug Rat Race," *Brookings Papers on Economic Activity*, 2010, *2010* (1), 129–176.

Rangel, Antonio, "Forward and Backward Intergenerational Goods: Why Is Social Security Good for the Environment?," *American Economic Review*, 2003, *93* (3), 813–834.

Razin, Assaf and Efraim Sadka, "Aging population: The complex effect of fiscal leakages on the politico-economic equilibrium," *European Journal of Political Economy*, 2007, *23* (2), 564–575.

_ , _ , and Phillip Swagel, "The Aging Population and the Size of the Welfare State," *Journal of Political Economy*, 2002, *110* (4), 900–918.

Reardon, Sean F., "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations," in Greg J. Duncan and Richard J. Murnane, eds., *Whither opportunity? Rising inequality, schools, and childrens life chances*, 2011, chapter 5, pp. 91–116.

Rentería, Elisenda, Guadalupe Souto, Iván Mejía-Guevara, and Concepció Patxot, "The Effect of Education on the Demographic Dividend," *Population and Development Review*, 2016, *42* (4), 651–671.

Restuccia, Diego and Carlos Urrutia, "Persistence of Earnings: Intergenerational The Role of Early and College Education," *The American Economic Review*, 2004, *94* (5), 1354–1378.

Rindfuss, Ronald R., Larry Bumpass, and Craig St. John, "Education and Fertility: Implications for the Roles Women Occupy," *American Sociological Review*, 1980, *45* (3), 431–447.

Ronald Lee and Andrew Mason, *Population Aging and the Generational Economy: A Global Perspective*, Cheltenham: Edward Elgar, 2011.

Roodman, David, "How to do xtabond2: An introduction to difference and system GMM in Stata," *Stata Journal*, 2009, *9* (1), 86–136.

Sánchez-Romero, Miguel, Gemma Abio, Concepció Patxot, and Guadalupe Souto, "Contribution of demography to economic growth," *SERIEs*, 2017.

Sardon, Jean-Paul, "Évolution Démographique Récente Des Pays Développés," *Population*, 2006, *61* (3), 197–266.

Shelton, Cameron A., "The aging population and the size of the welfare state: Is there a puzzle?," *Journal of Public Economics*, 2008, *92* (3-4), 647–651.

Soares, Jorge, "A dynamic general equilibrium analysis of the political economy of public education," *Journal of Population Economics*, jun 2006, *19* (2), 367–389.

Solon, Gary, "Intergenerational mobility in the labor market," in "Handbook of Labor Economics," Vol. 3 1999, pp. 1761–1800.

Stiglitz, J. E., "The demand for education in public and private school systems," *Journal of Public Economics*, 1974, *3* (4), 349–385.

Tabellini, Guido, "A Positive Theory of Social Security," *Scandinavian Journal of Economics*, 2000, *102* (3), 523–545.

Tepe, Markus and Pieter Vanhuysee, "Are Aging OECD Welfare States on the Path to Gerontocracy? Evidence from 18 Democracies, 1980-2002," *Journal of Public Policy*, 2009, *29* (1), 1–28.

Tolbert, Charles M. and Molly Sizer, "U.S. Commuting Zones and Labor Market Areas: A 1990 update," 1996.

Wooldridge, Jeffrey M., *Econometric analysis of cross section and panel data*, MIT Press, 2010.